

FIGURE 1

CCAGGTCCAACCTGCACCTCGGTTCTATCGATTGAATCCCCGGGGATCCTCTAGAGATCCCT
CGACCTCGACCCACGCGTCCGCCAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGA
CAGGCCAGGCAGGTGGGCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCTGAGGCCCCAGC
AAGGGCTAGGTTCCATCTCCAGTCCAGGACACAGCAGCGGCCACCATGGCCACGCTGGGC
TCCAGCAGCATCAGCAGCCCCCAGGACCGGGGAGGCACAGTGGCCCCCACCACCCGGAGG
AGCAGTCTCTGCCCTGTCTCGGGGATGACTGATTCTCTCCGCCAGGCCACCCAGAGGAGA
AGGCCACCCCGCTGGAGGCACAGGCCATGAGGGCTCTCAGGAGGTGCTGCTGATGTGGCT
TCTGGTGTGGCAGTGGGCGGCACAGAGCACGCTACCGGCCCGGCGTTAGGGTGTGTGCT
GTCCCGGGCTCACGGGGACCTGTCTCCGAGTCGTTCGTGCAGCGTGTGTACCAGCCCTTCC
TCACCACCTGCACGGGCACCGGGCCTGCAGCACCTACCGAACCATTTATAGGACCGCCTAC
CGCCGACGCCCTGGGCTGGCCCTGCCAGGCCTCGCTACGCGTGTGCCCCGGCTGGAAGAG
GACCAGCGGGCTTCTTGGGGCTGTGGAGCAGCAATATGCCAGCCGCCATGCCGAACGGAG
GGAGCTGTGTCCAGCCTGGCCGTGCCGTGCCCTGCAGGATGGCGGGGTGACACTTGGCAG
TCAGATGTGGATGAATGCAGTGTCTAGGAGGGGCGGCTGTCCCAGCGCTGCATCAACACCGC
CGGCAGTTACTGGTGCCAGTGTGGGAGGGGCACAGCCTGTCTGCAGACGGTACACTCTGTG
TGCCCAAGGGAGGGCCCCCAGGGTGGCCCCCAACCGACAGGAGTGGACAGTGAATGAAG
GAAGAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGTGGTGT
GGCCCCACTGCACAGCCTGGCCTCGCAGGCACTGGAGCATGGGCTCCCGGACCCCGGCAGCC
TCCTGGTGCCTCTTCCAGCAGCTCGGCCGCATCGACTCCCTGAGCGAGCAGATTTCCTTC
CTGGAGGAGCAGCTGGGGTCTGCTCCTGCAAGAAAGACTCGTGACTGCCAGCGCCCCAGG
CTGGACTGAGCCCCACGCGCCCTGCAGCCCCCATGCCCTGCCAACATGCTGGGGGTC
CAGAAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGCAGGCCTTCTCCTTTCTCTCTC
CCCTTCCCTCGGAGGGTCCCCAGACCCTGGCATGGGATGGGTGGGATTTTCTTGTGAAT
CCACCCTTGCTACCCCCACCTGGTTACCCCAACGGCATCCAAGGCAGGTGGGCCCTCA
GCTGAGGAAGGTACGAGTTCCCTGTCTGGAGCCTGGGACCATGGCACAGGCCAGGCAGCC
CGGAGGCTGGGTGGGGCTCAGTGGGGGTGCTGCCTGACCCCCAGCACAAATAAAATGAA
CGTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGACTCT
AGAGTCGACCTGCAGAAGCTTGGCCGCATGGCCCAACTTGTTATTGCAGCTTATAATGGT
TACAAAT

FIGURE 2

MTDSPPPGHPPEEKATPPGGTGHEGLSGGAADVASGVGSGRHRARLPARPLGCVLSRAHGDPV
SESFVQRVYQPFLLTTCDGHRACSTYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGAC
GAAICQPPCRNGGSCVQPGRCRCPAGWRGDTQCQSDVDECSARRGGCPQRCINTAGSYWCQCW
EGHSLSADGTLCPVKGPPRVAPNPTGVDSAMKEEVQRLQSRVDLLEEKQLVLAPLHSLAS
QALEHGLPDPGSLLVHSFQQLGRIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

amino acids 1-19

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 93-97, 270-274

N-myristoylation sites.

amino acids 19-25, 78-84, 97-103, 100-106, 103-109, 157-163,
191-197, 265-271

Amidation site.

amino acids 26-30

Aspartic acid and asparagine hydroxylation site.

amino acids 152-164

Cell attachment sequence.

amino acids 130-133

EGF-like domain cysteine pattern signature.

amino acids 123-135

FIGURE 3

CGCTCGCCCCGTGCCCCCTCGCCTCCCCGCAGAGTCCCCCTCGCGGCAGCAGATGTTGTGTGGG
GTCAGCCCCACGGCGGGGACAT**T**GGTGAAATTCCTCGCGCCTCAGCAGCTACTGGCCCCGTATC
CGGTTCTTGTGTGCCCTCGGGCATCACCAACATAGCCATCGATTCGGGGAGCAGGCCCTTGAA
CCGGGGCATTGCTGCTGTCAAGGAGGATGCAGTCGAGATGCTGGCCAGCTACGGGGCTGGCGT
ACTCCCTCATGAAGTTCTTCACGGGTCCCATGAGTGACTTCAAAATGTGGGCTTGGTGT
GTGAACAGCAAGAGAGACAGGACCAAAGCCGTCTGTGTATGGTGGTGGCGAGGGCCATCGC
TGCCGCTCTTTCACACACTGATAGCTTATAGTGATTTAGGATACTACATTATCAATAAACTGC
ACCATGTGGACGAGTCCGTGGGGAGCAAGACGAGAAGGGCTTCCTGTACCTCGCCGCTTT
CTTTTCATGGACCAATGGCATTGGACCCATGCTGGCATTCTCTTAAACACAAATACAGTTT
CCTGGTGGGATGTGCCTCAATCTCAGATGTCATAGCTCAGGTTGTTTTGTAGCCATTTTGC
TTCACAGTCACCTGGAATGCCGGGAGCCCTGCTCATCCCGATCCTCTCCTTGTACATGGGC
GCACTTGTGCGCTGCACCACCCTGTGCCCTGGGCTACTACAAGAACATTACGACATCATCCC
TGACAGAAAGTGGCCCCGAGCTGGGGGAGATGCAACAATAAGAAAGATGCTGAGCTTCTGGT
GGCCTTTGGCTCTAATTCTGGCCACACAGAGAATCAGTCGGCCTATTGTCAACCTCTTTGTT
TCCCCGGACCTTGGTGGCAGTTCTGCAGCCACAGAGGCAGTGGCGATTTTGACAGCCACATA
CCCCGTGGGTTCAGATGCCATACGGCTGGTTGACGGAATCCGTGCTGTGTATCCTGCTTTCG
ACAAGAATAACCCCGACAAACAACTGGTGAGCAGGACCAACAGCTCAGCGCAGCCACATC
AAGAAGTTACCTTCGTCTGCATGGCTCTGTCACTCAGCTCTGTTCCTGTGATGTTTTGGAC
ACCCAACGTGCTCGAGAAAATCTTGATAGACATCATCGGAGTGGACTTTGCCTTTGCAGAAC
TCTGTGTTGTTCTTCTGGGATCTTCTCCTTCTCCAGTTCAGTCACAGTGAGGGCGCAT
CTCACGGGTGGCTGATGACATCGAAGAAAACCTTCGTCTTGGCCCCAGCTCTGTGCTGGC
GATCATCGTCTCATCGCCAGCTCTGTGGTCTTACCCTACCTCGGGGTGCACGGTGGGACCC
TGGGCGTGGGCTCCCTCCTGGCGGGCTTTGTGGGAGAATCCACCATGGTCGCCATCGCTGGC
TGCTATGTCTACCGGAAGCAGAAAAGAAGATGGAGAATGAGTCGGCCACGGAGGGGGAAGA
CTCTGCCATGACAGACATGCGCTCCGACAGAGGAGGTGACAGACATCGTGGAAATGAGAGAGG
AGAATGAAT**T**AAGGCACGGGACGCCATGGGCATGTCAGGGACGGTCAGTCAGGATGACATTC
GGCATCATCTCTCCCTCTCCCATCGTATTTTGTTCCTTTTTTTTGTTTTGTTTTGGTAAT
GAAAGAGGCCTTGATTTAAAGGTTTCGTGTCAATTCTCTAGCATACTGGGTATGCTCACACT
GACGGGGGGACCTAGTGAATGGTCTTTACTGTGTCTATGTAATAACAAACGAAACAACTGAC
TTCATACCCCTGCCCTCAGAAAACCCAAAAGACACAGCTGCTCACGGTTGACGTTTGTGTCC
TCCTCCCCTGGACAATCTCCTCTTGAACCAAAGGACTGCAGCTGTGCCATCGCGCCTCGGT
CACCTGTCACAGCAGGCCACAGACTCTCCTGTCCCCCTTCATCGCTCTTAAGAATCAACAGG
TTAAACACTCGGCTTCCTTTGATTGTCTCCAGTCACATGGCCGTACAAAGAGATGGAGCCC
CGGTGGCCTCTTAAATTTCCCTTCTGCCACGGAGTTCGAACCATCTACTCCACACATGACG
GAGGCGGGTGGCACGCTGCAGCCCGAGTCCCCGTTACACTGAGGAACGGAGACCTGTGAC
CACAGCAGGCTGACAGATGGACAGAATCTCCCGTAGAAAGGTTTGGTTTGAAATGCCCGGG
GGCAGCAAACTGACATGGTTGAATGATAGCATTTCACTCTGCGTTCTCTAGATCTGAGCAA
GCTGTCAGTTCTCACCCCCCGGTGTATATACATGAGCTAATTTTTTAAATTTGTACAAAA
GCGCATCTCCAGATTCAGACCCTGCCGCATGACTTTTCTGAAGGCTTGCTTTTCCCTCGC
CTTTCTGAAGGTCGCATTTAGAGCGAGTCACATGGAGCATCTTAACCTTTGCATTTTAGTTTT
TACAGTGAACCTGAAGCTTTAAGTCTCATCCAGCATTCATGACAGGTTCTGCTAGGGTAGA
TTTTGAAGTAGATATATTACCTGGTTCGTCTATCCTTAGTCATAACTCTGCGGTACAGGTA
TTGAGATGTACTACGGTACTTCCCTCCCACACCATAGATAAGCAAGACATTTTATAACG
ATACAGAGTCACTATGTGGTCTCCTCCCTGAAATAACGCATTCGAATCCATGCAGTCAGTA
TATTTTTCTAAGTTTGGGAAAGCAGGTTTTTTCTTTAAAAAAATATAGACACGGTTTCACT
AAATGATTATTAGTCAGAAATCTTAGACTGAAAGAACCTAAACAAAAAAATATTTTAAGATA
TAAATATATGCTGTATATGTTATGTAATTTATTTTAGGCTATAATACATTTTCTATTTTTCG
ATTTTCAATAAATGTCTCTAATACAAAAA

FIGURE 4

MVKFPALTHYWPLIRFLVPLGITNIAIDFGEQALNRGIAAVKEDAVEMLASYGLAYSLMKFF
TGPMSEDFKNVGLVFVNSKRDRTKAVLCMVVAGAIAAVFHTLIAYSDLGYIINKLHHVDESV
GSKTRRAFLYLAAPFMDAMAWTHAGILLKHKYSFLVGCASISDVIAQVVFVAILLHSHLEC
REPLLIPIILSLYMGALVRCTTLC LGYYKNIHDIIPDRSGPELGGDATIRKMLSFWWPLALIL
ATQRI SRPIVNL FVS RDLGGSSAATEAVAILTATYPVGHMPYGWLTEIRAVYPAFDKNPNPSN
KLVSTSNVTAAHIKKFTFVCMALSLTLCFVMFWTPNVSEKILIDIIGVDFAFABELCVVPLR
IFSFFPVPVTVRAHLTGWLMTLKKTFVLAPSSVLRIIVLIASLVVLPYLGVBHATLGVGSL
AGFVGESTMVAIAACYVYRKQKKMENESATEGEDSAMTDMPTTEEVTDIVEMREENE

Transmembrane domains:

amino acids 86-106, 163-179, 191-205, 237-253, 327-343, 357-374,
408-423, 431-445

FIGURE 5

CCTGACAGAAGTGCCCCGGAGCTGGGGGAGATNCAACATTAAGAAGATGCTGAGCTTCTGGT
GCCNTTTGGCTCTAATTCTGGCCACACAGAGAANCAGTCGGCCTATTGTCAACCTCTTTGTT
TCCCGGGACCTTGGTGGCAGTTCTGCAGCCACAGAGGCAGTGGCGATTTTGACAGCCACATA
CCCTGTGGGTCACATGCCATACGGCTGGTTGACGGAAATCCGTGCTGTGTATCCTGCTTTCG
ACAAGAATAACCCAGCAACAACTGGTGAGCACGAGCAACACAGTCACGGCGGCCACATC
AAGAAGTTCACCTTCGTCTGCATGGCTCTGTCACTCACGCTCTGTTTCGTGATGTTTTGGAC
ACCCAACGTGTCTGNGAAAATCTTGATAGACATCATCGAGTGGACTTGCCTTTGCAGAAC
TCTGTGTTGTTCTTTGCGGATCTTCTCCTTCTTCCAGTTCAGTCAAGTGAGGGCGCAT
CTCACGGGTGGCTGATGACACTGAAGAAAACCTTCGTC

1000
900
800
700
600
500
400
300
200
100
0
1000
900
800
700
600
500
400
300
200
100
0
1000
900
800
700
600
500
400
300
200
100
0

FIGURE 6

TGACGGAATCCCGGGCTGGGTATCCTGGTTTNGACAAGATAAAACCCAGCAANAAATTGGG
GAGCAGGGCAAAACAGTNACGGGCAGCCACATCAAGAAGTTCACCTTNGTTTGNATGGNTC
TGTCAACTCACGCTNTGTTTCGTGATGTTTTGGACACCCAAAGTGTGAGAAAAATTTGAT
AGACATNATCGGAGTGGANTTTGCCTTTGCAGAANTTTGNGNTGTTCCCTTGCGGATTTTCT
CCTTTTTCCAGTTCAGTCACAGNGAGGGCGCATCTACCGGNGGNTGATGACANTGAAG
AAACCTTTGTCCTTGCCCCAGCTNTTTGGTGCGGATCATTGTCCTNATNGCCAGCCTTGT
GGTCCTACCTACCTGGGGGTGCACGGTGCGACCCTGGGCGTGGGTCCCTCCTGGCGGGCA

FIGURE 7

TATTCCCAGTTCCGGTCACGGGGAGGGCGCATNTCACCGGGTGGCTGANGACACTGAAGAAA
ACCTTNGTCCTTGCCCCCAGNTTTGTGNTGCGGATNATCGTCCTCATCGCCAGCCTNGTGGT
CCTACCCTACCTGGGGGTGCACGGTGAGAC

5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20

FIGURE 8

GCCCCGCGCCCGCGCCGGGCGCCCGAAGCCGGGAGCCACCGCCATGGGGGCCCTGCCTGGGA
GCCTGCTCCCTGCTCAGCTGCGCGTCTGCCTCTGCGGCTCTGCCCCCTGCATCCTGTGCAG
CTGCTGCCCCGCCAGCCGCAACTCCACCGTGAGCCGCCTCATCTTCAAGTTCTTCTCTTCC
TGGGGGTGCTGGTGTCCATCATTATGCTGAGCCCGGGCGTGGAGAGTCACTCTACAAGCTG
CCCTGGGTGTGTAGGAGGGGGCCGGGATCCCCACCCTCTCTGAGGGCCACATCGACTGTGG
CTCCCTGCTTGCTTACCGCGCTGTCTACCGCATGTGCTTCGCGACGGCGGCCCTTCTTCTTCT
TCTTTTTCACCCCTGCTCATGCTCTGCGTGAGCAGCAGCCGGGACCCCGGGCTGCCATCCAG
AATGGGTTTTGGTTCTTTAAGTTCTTGATCCTGGTGGGCCCTACCGTGGGTGCCTTCTACAT
CCCTGACGGCTCCTTCACCAACATCTGGTTCTACTTCGGCGCTCGTGGGCTCCTTCCTCTTCA
TCCTCATCCAGCTGGTGTGCTCATCGACTTTGCGCACTCCTGGAACAGCGGTGGTGGGC
AAGGCCGAGGAGTGCATTCCCGTGCCTGGTACGCAGGCCCTTCTTCTTCTCACTCTCCTCTT
CTACTTGTCTGATCGCGGCCGTGGCGCTGATGTTTCATGTACTACACTGAGCCAGCGGCT
GCCACGAGGGCAAGGTCTTCATCAGCCTCAACCTCACCTTCTGTGTCTGCGTGTCCATCGCT
GCTGTCTGCCCAAGGTCCAGGACGCCAGCCAACTCGGGTCTGCTGCAAGCCCTCGGTTCAT
CACCTCTACACCATGTTTGTACCTGGTCAGCCCTATCCAGTATCCCTGAACAGAAATGCA
ACCCCATTTGCCAACCCAGCTGGGCAACGAGACAGTTGTGGCAGGCCCCGAGGGCTATGAG
ACCCAGTGGTGGGATGCCCCGAGCATTTGTGGGCTCATCATCTTCTCCTGTGCACCTCTT
CATCAGTCTGCGCTCCTCAGACCACCGGCAGGTGAACAGCCTGATGCAGACCGAGGAGTGCC
CACCTATGTAGACGCCACACAGCAGCAGCAGCAGCAGGTGGCAGCCTGTGAGGGCCGGGCC
TTTGACAACGAGCAGGACGGCGTCACTTACAGCTACTCCTTCTTCCATTCTGCTGGTGTCT
GGCCTCACTGCACGTATGATGACGCTCACCACTGGTACAAGCCCGGTGAGACCCGAAGA
TGATCAGCACGTGGACCGCCGTGTGGGTGAAGATCTGTGCCAGCTGGGCAGGGCTGCTCCTC
TACCTGTGGACCTGTAGTACCCCACTCCTCCTGCGCAACCGCGACTTCAGCTTGAGGCAGCCT
CACAGCCTGCCATCTGGTGCTCCTGCCACCTGGTGCCTCTCGGCTCGGTGACAGCCAACTT
GCCCCCTCCCCACACCAATCAGCCAGGCTGAGCCCCCACCCTGCCCCAGCTCCAGGACCTG
CCCCTGAGCCGGGCCTTCTAGTCGTAGTGCTTCAGGGTCCGAGGAGCATCAGGCTCCTGCA
GAGCCCCATCCCCGCCACACCCACACGGTGGAGCTGCCTCTTCTTCCCCCTCCTCCCTGT
TGCCCATACTCAGCATCTCGGATGAAAGGGCTCCCTTGTCCTCAGGCTCCACGGGAGCGGGG
CTGCTGGAGAGAGCGGGGAATCCCAACACAGTGGGGCATCCGGCACTGAAGCCCTGGTGT
CCTGGTCACTCCCCAGGGGACCTGCCCTTCTCTGAGACTTCGTGCCTTACTGAGTCTCT
AAGACTTTTTCTAATAAACAGCCAGTGCCTGTAAAAAAA

FIGURE 9

MGACLGACSLSCASCILCSGPCASRNSTVSRLIPTFFFLGVLVSIIMLSPGVE
SPLYKLPWVCEEGAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFFFTLLMLCVSSSRD
PRAAIQNGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSLFLILQLVLLIDFAHSW
NQRWLKGAEECDRAWYAGLFFFTLLFYLLSIAAVALMFMYTTEPSGCHEGKVFISLNLTPC
VCVSIAAVLPKVQDAQPNSGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVA
GPEGYETQWWDAPSIVGLIIFLLCTLFISLRSSDHRQVNSLMQTECCPPMLDATQQQQQVA
ACEGRAFDNEQDGVITYSYFFHFCLVLAHLVMMTLTNWYKPGETRMISTWTAVVWKICAS
WAGLLLYLWTLVAPLLLRNRDFS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

FIGURE 10

GAGCGAGGCCGGGACTGAAGGTGTGGGTGTTCGAGCCCTCTGGCAGAGGGTTAACCTGGGTC
AAATGCACGGATTCTCACCTCGTACAGTTACGCTCTCCGCGGCACGTCGCGAGGACTTGA
AGTCTTGAGCGCTCAAGTTTGTCCGTAGGTTCGAGAGAAGGCCATGGAGGTGCCGCCACCGGC
ACCGCGGAGCTTTCTCTGTAGAGCATTGTGCCTATTTCCCGAGTCTTTGCTGCCGAAGCTG
TGACTGCCGATTTCGGAAGTCCTTGAGGAGCGTCAGAAGCGGCTTCCCTACGTCCCAGAGCCC
TATTACCCGGAATCTGGATGGGACCGCCTCCGGGAGCTGTTTGCGCAAAGATGAACAGCAGAG
AATTTCAAAGGACCTTGCTAATATCTGTAAGACGGCAGCTACAGCAGGCATCATTGGCTGGG
TGTATGGGGGAATACCAGCTTTTATTTCATGCTAAACAACAATACATTGAGCAGAGCCAGGCA
GAAATTTATCATAACCGGTTTGATGCTGTGCAATCTGCACATCGTGCTGCCACACGAGGCTT
CATTCTGTTATGGCTGGCGCTGGGGTTGGAGAACTGCAGTGTTTGTGACTATATTCAACACAG
TGAACACTAGTCTGAATGTATACCGAAATAAAGATGCCTTAAGCCATTTTGTAAATTGCAGGA
GCTGTACCGGAAGTCTTTTTAGGATAAAACGTAGGCCTGCGTGGCCTGGTGGCTGGTGGCAT
AATTGGAGCCTTGCTGGGCACTCCTGTAGGAGGCCTGCTGATGGCATTTCAGAAGTACGCTG
GTGAGACTGTTTCAGGAAAGAAAAACAGAAGGATCGAAAGGCACTCCATGAGCTAAAACTGGAA
GAGTGGAAAGGCAGACTACAAGTTACTGAGCACCTCCCTGAGAAAATTGAAAGTAGTTTACG
GGAAGATGAACCTGAGAATGATGCTAAGAAAAATTGAAGCACTGCTAAACCTTCCTAGAAACC
CTTCAGTAATAGATAAAACAAGACAAGGACTGAAAGTGCTCTGAACCTTGAAACTCACTGGAGA
GCTGAAGGGAGCTGCCATGTCCGATGAATGCCAACAGACAGGCCACTCTTTGGTCAGCCCTGC
TGACAAAATTTAAGTGCTGGTACCTGTGGTGGCAGTGCGCTTGTCTTTTCTTTTCTT
TTTAACTAAGAATGGGCTGTGTACTCTCACTTTACTTATCCTTAAATTTAAATACATACT
TATGTTTGTATTAATCTATCAATATATGCATACATGGATATATCCACCCACCTAGATTTTAA
GCAGTAAATAAAACATTTTCGAAAAGATTAAAGTTGAATTTTACAGTTT

FIGURE 11

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA23318
><subunit 1 of 1, 285 aa, 1 stop
><MW: 32190, pI: 9.03, NX(S/T): 2
MEVPPPPAPRSFLCRALCLFPRVFAAEAVTADSEVLEERQKRLPYVPEPYYPESGWDRLRELFL
GKDEQQRISKDLANICKTAATAGIIGWVYGGIPAFIHAKQQYIEQSQAIEYHNRFDAVQSAH
RAATRGFIRYGWRWGVRTAVFVTIFNTVNTSLNVYRNKDLSHFVIAGAVTGSLFRINVGLR
GLVAGGIIIGALLGTPVGGLLMAFQKYAGETVQERKQKDRKALHELKLEEWKGRLQVTEHLPE
KIESSLREDEPENDAKKIEALLNLPRNPVIDKQDKD
```

Important Features:

Signal Peptide:

amino acids 1-24

Transmembrane domains:

amino acids 76-96 and 171-195

N-glycosylation site:

amino acids 153-156

FIGURE 12

CGGAAGTCCCTTGAGGAGCGTCAGAAGCGGCTTCCCTACGTCCCAGAGCCCTATTACCCGGA
ATCTGGATGGGACCGCTCCGGGAGCTGTTTGGCAAAGATGAACAGCAGAGAATTTCAAAGGA
CCTTGCTAATATCTGTAAGACGGCAGCTACAGCAGGCATCATTGGCTGGGTGTATGGGGGAA
TACCAGCTTTTATTCATGCTAAACAACAATACATTGAGCAGAGCCAGGCAGAAATTTATCAT
AACCGGTTTGATGCTGTGCAATCTGCACATCGTGCTGCCACACGAGGCTTCATTGTTTCATG
GCTGGCGCCGAACC

FIGURE 13

TCAAGTTTGTCCGTAGGTCGAGAGAAGGCCATGGAGGTGCCGCCACCGGCACCGCGGAGCTT
TTTTCTGTAGAGCATTGTGCCTATTTCCCCGAGTTTTTGCTGCCGAAGCTGTGACTGCCGAT
TCGGAAGTCCTTGAGGAGCGTCAGAAGCGGCTTCCCTACGTCCCAGAGCCCTATTACCCGGA
ATTTGGATGGGACCGCCTCCGGGAGCTGTTTGGCAAAGATGAACAGCAGAGAATTTCAAAGG
ACCTTGCTGATATNTGTAAGACGGCAGCTACAGCAGGCATCATTGGCTGGGTGTATGGGGGA
ATACCAGCTTTTATTCATGNTAAACAACAATACATTGAGCAGAGCCAGGCAGAAATTTATNA
TAACC

100
90
80
70
60
50
40
30
20
10
0
A
T
C
G
A
T
C
G
T
A
G
G
T
C
G
A
G
A
A
G
G
C
C
A
T
G
G
A
G
G
T
G
C
C
G
C
C
A
C
C
G
G
C
A
C
C
G
C
G
G
A
G
C
T
T
T
T
T
C
T
G
T
A
G
A
G
C
A
T
T
G
T
G
C
C
T
A
T
T
T
C
C
C
C
G
A
G
T
T
T
T
T
G
C
T
G
C
C
G
A
A
G
C
T
G
T
G
A
C
T
G
C
C
G
A
T
T
C
G
G
A
A
G
T
C
C
T
T
G
A
G
G
A
G
C
G
T
C
A
G
A
A
G
C
G
G
C
T
T
C
C
C
T
A
C
G
T
C
C
C
A
G
A
G
C
C
T
A
T
T
A
C
C
C
G
G
A
A
T
T
T
G
G
A
T
G
G
G
A
C
C
G
C
C
T
C
C
G
G
G
A
G
C
T
G
T
T
T
G
G
C
A
A
A
G
A
T
G
A
A
C
A
G
C
A
G
A
G
A
A
T
T
T
C
A
A
A
G
G
A
C
C
T
T
G
C
T
G
A
T
A
T
N
T
G
T
A
A
G
A
C
G
G
C
A
G
C
T
A
C
A
G
C
A
G
G
C
A
T
C
A
T
T
G
G
C
T
G
G
G
T
G
T
A
T
G
G
G
G
G
A
A
T
A
C
C
A
G
C
T
T
T
T
A
T
T
C
A
T
G
N
T
A
A
C
A
C
A
A
T
A
C
A
T
T
G
A
G
C
A
G
A
G
C
C
A
G
G
C
A
G
A
A
A
T
T
A
T
N
A
T
A
A
C
C

FIGURE 14

GAGCCGCCGCCGCGCGCGCGCCGCGCACTGCAGCCCCAGGCCCGGCCCCCCACCCACGTCT
GCGTTGTGCCCCGCCCTGGGCCAGGCCCCAAAGGCAAGGACAAAGCAGCTGTCAGGGAACCT
CCGCCGAGTCGAATTTACGTGCAGCTGCCGGCAACCACAGGTTCCAAGATGCTTTGCGGGG
GCTTCGCGTGTTC AAGAACTGCCTGTGCGCCCTCAACCTGCTTTACACCTTGGTTAGTCTG
CTGCTAATTGGAATTGCTGCGTGGGGCATTGGCTTCGGGCTGATTTCCAGTCTCCGAGTGGT
CGGCGTGGTCATTGCAGTGGGCATCTTCTGTTCCTGATTGCTTTAGTGGGTCTGATTGGAG
CTGTAAAACATCATCAGGTGTGCTATTTTTTATATGATTATTCTGTTACTTGTATTTATT
GTTCAGTTTTCTGTATCTTGCCTTGTTTAGCCCTGAACCAGGAGCAACAGGGTCAGCTTCT
GGAGGTTGGTTGGAACAATACGGCAAGTGCTCGAAATGACATCCAGAGAAATCTAAACTGCT
GTGGGTTCCGAAGTGTAAACCAAATGACACCTGTCTGGCTAGCTGTGTTAAAGTGACCAC
TCGTGCTCGCCATGTGCTCCAATCATAGGAGAATATGCTGGAGAGGTTTTGAGATTTGTTGG
TGGCATTGGCCGTGTTCTTCAGTTTTACAGAGATCCTGGGTGTTTGGCTGACCTACAGATACA
GGAACCAGAAAGACCCCGCGCGCAATCTAGTGCAATCTCTT TGAT GAGAAAAACAAGGAAGAT
TTCCTTTCGTATTATGATCTTGTTCACTTCTGTAATTTTCTGTTAAGCTCCATTTGCCAGT
TTAAGGAAGGAACACATATCTGGAAGTACCTTATTGATAGTGAATTATATATTTTTACT
CTATGTTTCTCTACATGTTTTTTCTTTCCGTTGCTGAAAAATATTTGAAACTGTGGTCTC
TGAAGCTCGGTGGCACCTGGAATTTACTGTATTATTGTGCGGCACCTGTCCACTGTGGCCTT
TCTTAGCATTTTTACCTGCAGAAAAACTTGTATGTTAGTACCCTGTGTTGGTTATATGGTGAA
TCTGAACGTACATCTCACTGGTATAATTATATGTAGCACTGTGCTGTGTAGATAGTTCCCTAC
TGAAAAAAGAGTGGAATTTATTAATAATCAGAAAGTATGAGATCCTGTTATGTTAAGGAAAA
TCCAAATTCCCAATTTTTTTTGGTCTTTTAGGAAAGATTGTTGTTGTTAAAAAGTGTTAGTA
TAAAAATGATAATTTACTTGTAGTCTTTTATGATTACACCAATGTATTCTAGAAATAGTTAT
GTCCTTAGGAAATTGTGGTTTAATTTTTGACTTTTACAGGTAAGTGCAAAAGGAGAAGTGTTTT
CATGAAATGTTCTAATGTATAATAACATTACCTTCAGCCTCCATCAGAAATGGAACGAGTTT
TGAGTAATCAGGAAGTATATCTATATGATCTTGATATTGTTTTATAATAATTTGAAGTCTAA
AAGACTGCATTTTTAAACAAGTTAGTATTAATGCGTTGGCCCCACGTAGCAAAAAGATATTTG
ATTATCTTAAAAATTGTTAAATACCGTTTTTCATGAAATTTCTCAGTATTGTAACAGCAACTT
GTCAAACCTAAGCATATTTGAATATGATCTCCCATAAATTTGAAATGAAATCGTATTGTGTG
GCTCTGTATATTCTGTTAAAAAATTAAAGGACAGAAACCTTTCTTTGTGTATGCATGTTTGA
ATTAAAAGAAAGTAATGGAAG

FIGURE 15

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA39979
><subunit 1 of 1, 204 aa, 1 stop
><MW: 22147, pI: 8.37, NX(S/T): 3
MVCGGFACSKNCLCALNLLYTLVSLLLIGIAAWGIGFGLISSLRVVGVIAGVIFLFLIALV
GLIGAVKHHQVLLFFYMIILLVFIQFSVSCACLALNQEQQQLLEVGNNTASARNDIQR
NLNCCGFRSVNPNDTCLASCVKSDHSCSPCAPIIGEYAGEVLRFRVGGIGLFFSFTEILGVWL
TYRYRNQKDPKANPSAFL
```

Signal Peptide:

amino acids 1-34

Transmembrane domains:

amino acids 47-63, 72-95 and 162-182

FIGURE 16

TGATTGGAGCTGTAAAAAANTCTTCAGGTGTTGTNATTTTTTTATATGATTATTCTGTAANT
TGTATTTATTGTTTCAGTTTTNTGTATCTTGCGCTTGTTTAGCCNTGAACCAGGAGCAACAGG
GTCAGNTTNTGGAGGTGGGTGGAACAATACGGCAAGTGCTCGAAATGACATCCAGAGAAAT
NTAAACTGCTGTGGGTTCGGAAGTGTTAACCCAAATGACACCTGINTGGCTAGCTGTGTTAA
AAGTGACCACTNGTGCTCGCCATGTGCTCCAATCATAGGAGAATATGCTGGAGAGGTTTTGA
GATTGTGTTGGTGCCATTGGCCTGTNTTCAGTTTTACAGAGATCCTGGGTGTTTGGCTGACC
TACAGATACAGGAACCAG

100
90
80
70
60
50
40
30
20
10
0
100
90
80
70
60
50
40
30
20
10
0

[illegible]

AATCCCAAAATCCCAATTTTTTGGNCTTTTAGGGAAGATGTGTGTGGTAAAAAGTGT
TAGTATAAAAATGATAATTTACTTGTAGTCTTTTATGATTACACCAATGTATTCTAGAATAG
TTATGTCTTAGGAAATGTGTGTTAATTTTTGACTTTTACAGGTAAGTGC AAGGAGAAGTG
GTTTCATGAAATGTTCTAATGTATAATAACATTTACCTTCAGCTCCCATCAGAATGGAACG
AGTTTGTAGTAATCCAGGAAGTATATCTATATGATCTTGATATTGTTTTATATAATTTGAAG
TCTAAAAGACTGCATTTTAAACAAGTTAGTATTAATGCGTTGGCCACGTAGCAAAAAGAT
ATTTGATTATCTTAAAAAATGTTAAATACCGTTTTTCATGAAAGTTCTCAGTATTGTAACAGC
AACTTGTCAAACCTAAGCATATTTGAATATGATCTCCCATAAATTTGAAATTGAAATCGTATT
GTGTGGAGGAAATGGCAATCTTATGTGTGCTGAAGGACACAGTAAGAGCACCAAGTTGTGCC
CCACTTGC

[illegible]

ATGATTATTCTGTTACTTGTATTTATTGTTCAAGTTTATGGTATCTTGCAGCTTGTTAGCCC
CTGA AACAGGAGCAACAGGGNNCAGCTTCTCGAGGTTGGTTGGCAACATCAGGGCCAAG
TGACTCCGCAAATGACATCCAGAGAAATCCTAAACTGCTGTGGGTTCCGAAGTGTTAAACC
AAATGACACCTGTCTGGCTNGCTGTGTTAAAGTGACCACTCGTGCTGCCATGTGCTCCA
CTATAGGAGGAATATGC

FIGURE 19

CAGTCACCATGAAGCTGGGCTGTGTCTCATGGCCTGGGCCCTCTACCTTTCCCTTGGTGTG
CTCTGGGTGGCCCCAGATGCTACTGGCTGCCAGTTTTGAGACGCTGCAGTGTGAGGGACCTGT
CTGCACCTGAGGAGAGCAGCTGCCACACGGAGGATGACTTGACTGATGCAAGGGAAGCTGGCT
TCCAGGTCAAGGCTTACACTTTTCAGTGAACCTTCCACCTGATTGTGTCTATGACTGGCTG
ATCTCCAAGGTCCAGCCCAAGCCAGTTTTTGAAGGGGACCTGCTGGTTCTGCGCTGCCAGGC
CTGGCAAGACTGGCCACTGACTCAGGTGACCTTCTACCGAGATGGCTCAGCTCTGGGTCCCC
CCGGGCCCTAACAGGGAATTCTCCATCACCGTGGTACAAAAGGCAGACAGCGGGCATTACCA
TGCAGTGGCATCTTCCAGAGCCCCGGTCTGGGATCCCAGAAAAGCATCTGTTGTGGCTAT
CACAGTCCAAAGAACTTTTCCAGCGCCAATTCTCAGAGCTGTACCCCTCAGCTGAACCCCAAG
CAGGAAGCCCCATGACCTGAGTTGTGACACAAAGTTGCCCTCAGAGGTCAGCTGCCCGC
CTCCTCTTCTCCTTCTACAAGGATGGAAGGATAGTGCAAAGCAGGGGGCTCTCCTCAGAATT
CCAGATCCCCACAGCTTCAGAAGATCACTCCGGGTCACTACTGGTGTGAGGCAGCCACTGAGG
ACAACCAAGTTTGAAACAGAGCCCCAGCTAGAGATCAGAGTGCAGGGTGCTTCCAGCTCT
GCTGCACCTCCACATTGAATCCAGCTCCTCAGAAATCAGCTGCTCCAGGAATGCTCCTGA
GGAGGCCCTGGGCTCTGCTCCGCCCAACCCATCTTCTGAGGATCCAGGCTTTTCTT
CTCCTCTGGGGATGCCAGATCCTCATCTGTATCACCAGATGGGGCTTCTTCTCAAAACATG
CAGGATGTGAGAGTCTCCTCGGTCACTGCTCATGGAGTTGAGGGAATTATCTGGCCACCA
GAAGCCTGGGACCACAAAGGCTACTGCTGAATAGAAAGTAAACAGTTTATCCATGATCTCACT
TAACCAACCCCAATAAATCTGATTCTTTATTTCTCTTCTGTCTGCACATATGCATAAGTA
CTTTTACAAGTTGTCCCAGTGTTTTGTTAGAAATAATGTAGTTAGGTGAGTGTAATAAATTT
ATATAAAGTGAGAATTAGAGTTTAGCTATAAATTGTGTATTCTCTTAACACAAACAGAATTC
TGCTGTCTAGATCAGGAATTTCTATCTGTTATATCGACCAGAATGTTGTGATTTAAAGAGAA
CTAATGGAAGTGGATTGAATACAGCAGTCTCAACTGGGGCAATTTTGCCCCCAGAGGACA
TTGGGCAATGTTTGGAGACATTTTGGTCATTATCTTGGGGGGTTGGGGGATGGTGGGATGT
GTGCTACTGGCATCCAGTAAATAGAAGCCAGGGGTGCCGCTAAACATCCTATAATGCACAG
GGCAGTACCCACAAACGAAAAATAATCTGGCCAAATGTGAGTTGTACTGAGTTTGAAGAA
CCCCAGCCTAATGAACCCCTAGGTGTTGGGCTCTGGAATGGGACTTTGTCCCTTCTAATTAT
TATCTCTTTCCAGCCTCATTGAGCTATTCTTACTGACATACCACTCTTAGCTGGTGCTATG
GTCTGTTCTTTAGTTCTAGTTTGTATCCCTCAAAGCCATTATGTTGAAATCCTAATCCCC
AAGGTGATGGCATTAAAGAAGTGGGCCTTTGGGAAGTGATTAGATCAGGAGTGCAGAGCCCTC
ATGATTAGGATTAGTGCCCTTATTTAAAAAGGCCCCAGAGAGCTAACTACCCTTCCACCAT
ATGAGGACGTGGCAAGAAGATGACATGTATGAGAACCAAAAACAGCTGTGCGCAACACCG
ACTCTGTCGTTGCTTTGATCTTGAACCTTCCAGCCTCCAGAACTATGAGAAATAAAATCTCG
TTGTTTGTAGCCTAA

FIGURE 20

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA40594
><subunit 1 of 1, 359 aa, 1 stop
><MW: 38899, pI: 5.21, NX(S/T): 0
MKLGCVLMAWALYLSLGVLWVAQMLLAASFETLQCEGFPVCTEESSCHTEDDLTDAREAGFQV
KAYTFSEPFHLIVSYDWLILQGPAKPVFEGDLLVLRQCQAWQDWPLTQVTFYRDGSALGPPGP
NREFSITVVQKADSGHYHCSGIFQSPGPGIPETASVVAITVQELFPAPILRAVPSAEPQAGS
PMTLSCQTKLPLQRSARLLFSFYKDGRIVQSRGLSSEFQIPTASEDHSGSYWCEAATEDNQ
VWKQSPQLEIRVQGAASSAAPTTLNPAPQKSAAPGTAPEEAPGPLPPPPPTPSSDPGFSSPL
GMPDPHLYHQMGLLLKHMQDVRVLLGHLLMELRELSGHQKPGTTKATAE
```

Signal sequence:

amino acids 1-17

Leucine zipper pattern sequence:

amino acids 12-33

Protein kinase C phosphorylation site:

amino acids 353-355

FIGURE 21

CCCACGCGTCCGCCACGCGTCCGCCACGGGTCCGCCACGCGTCCGGGCCACCAGAAGTT
TGAGCCTCTTTGGTAGCAGGAGGCTGGAAGAAAGGACAGAAGTAGCTCTGGCTGTGATGGGG
ATCTTACTGGCCCTGCTACTCTCTGGGCCACCTAACAGTGGACACTATGGCCGTCCCATCCT
GGAAGTGCCAGAGAGTGTAAACAGGACCTTGGAAGGGGATGTGAATCTTCCCTGCACCTATG
ACCCCTGCAAGGCTACACCCAAGTCTTGGTGAAGTGGCTGGTACAAAGTGGCTCAGACCT
GTCACCATCTTTTCTACGTGACTCTTCTGGAGACCATATCCAGCAGGCAAAAGTACCAGGGCCG
CCTGCATGTGAGCCACAGGTTCCAGGAGATGTATCCCTCCAATTGAGCACCTGGAGATGG
ATGACCGGAGCCACTACACGTGTGAAGTCACCTGGCAGACTCTTGATGGCAACCAAGTCTGTG
AGAGATAAGATTACTGAGCTCCCGTGTCCAGAACTCTGTCTCCAAGCCACAGTGACAAAC
TGGCAGCGGTTATGGCTTCACGGTGCCCCAGGGAATGAGGATTAGCCTTCAATGCCAGGCTC
GGGGTTCTCTCCCATCAGTTATATTTGGTATAAGCAACAGACTAATAACCAAGGAACCCATC
AAAGTAGCAACCTAAGTACCTTACTCTTCAAGCCTGCGGTGATAGCCGACTCAGGCTCCTA
TTTCTGCACTGCCAAGGCCAGGTTGGCTCTGAGCAGCACAGCGACATTGTGAAGTTTGTGG
TCAAAGACTCCTCAAAGCTACTCAAGACCAAGACTGAGGCACCTACAACCATGACATACCCC
TTGAAGCAACATCTACAGTGAAGCAGTCTCTGGACTGGACCACTGACATGGATGGCTACCT
TGGAGAGACCAGTGCTGGGCCAGGAAAGAGCCTGCCTGTCTTTGCCATCATCCTCATCATCT
CCTTGTGCTGTATGGTGGTTTTTACCATGGCCTATATCATGCTCTGTGCGAAGACATCCCAA
CAAGAGCATGTCTACGAAGCAGCCAGGTAAGAAAGTCTCTCTCTTCCATTTTGAACCCGCT
CCCTGCCCTCAATTTTGATTACTGGCAGGAAATGTGGAGGAAGGGGGTGTGGCACAGACCC
AATCTTAAGGCCGGAGGCCCTTCAGGGTCAGGACATAGCTGCCTTCCCTCTCTCAGGCACCTT
CTGAGGTTGTTTTGGCCCTCTGAACACAAAGGATAAATTAGATCCATCTGCCTTCTGCTTCC
AGAATCCCTGGGTGGTAGGATCCTGATAATTAATTGGCAAGAATTGAGGCAGAAGGGTGGGA
AACCAGGACCACAGCCCCAAGTCCCTTCTTATGGGTGGTGGGCTCTTGGGCCATAGGGACA
TGCCAGAGAGGCCAACGACTCTGGAGAAACATGAGGGTGGCCATCTTCGCAAGTGGCTGCT
CCAGTGTATGAGCCAACTTCCAGAATCTGGGCAACAACACTCTGATGAGCCCTGCATAGGA
CAGGAGTACCAGATCATCGCCAGATCAATGGCAACTACGCCCGCCTGCTGGACACAGTTCC
TCTGGATTATGAGTTTCTGGCCACTGAGGGCAAAAGTGTCTGTTAAAAATGCCCATTAGGC
CAGGATCTGTGACATAAATTGCCTAGTCAGTCTTGCCTTGCATGGCCTCTTCCCTGCT
ACCTCTCTTCTGGATAGCCCAAAGTGTCCGCCCTACCAACACTGGAGCCGCTGGGAGTCACT
GGCTTTGCCCTGGAATTTGCCAGATGCATCTCAAGTAAGCCAGCTGTGGAATTTGGCTCTGG
GCCCTTCTAGTATCTCTGCCGGGGGCTTCTGTACTCCTCTCTAAATACCAGAGGGAAGATG
CCCATAGCACTAGGACTTGGTCATCATGCTACAGACACTATTCAACTTTGGCATCTTGCCA
CCAGAAGACCCGAGGGAGGCTCAGCTCTGCCAGCTCAGAGGACCAGCTATATCCAGGATCAT
TTCTCTTTCTCAGGGCCAGACAGCTTTTAATTGAAATTGTATTTCACAGGCCAGGGTTCA
GTTCTGCTCCTCCACTATAAGTCTAATGTTCTGACTCTCTCCTGGTGTCAATAAATATCTA
ATCATAACAGC

FIGURE 22

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45416
><subunit 1 of 1, 321 aa, 1 stop
><MW: 35544, pI: 8.51, NX(S/T): 0
MGILLGLLLLGHLTVDTYGRPILEVPESVTGPWKGDVNLPCYDPLQGYTQVLVKWLVRGS
DPVTIFLRDSSGDHIQQAKYQGRLVSHKVPDVSLLQLSTLEMDDRSHYTCEVTWQTPDGNQ
VVRDKITELRVQKLSVSKPTVTTGSGYGFTVPQGMRIQLQCQARGSPPISYIWKQQTNNQE
PIKVATLSTLLFKPAVIADSGSYFCTAKGVGSEQHSDIVKVVVDSSKLLKTKTEAPTTMT
YPLKATSTVKQSWDWTMDGYLGETSAGPGKSLPVFAIILIIISLCCMVVFTMAYIMLCRKT
SQEHVVEAAR

Signal Sequence:

amino acids 1-19

Glycosaminoglycan attachment site:

amino acids 149-152

Transmembrane domain:

amino acids 282-300

FIGURE 23

GCGCCGGGAGCCCATCTGCCCCAGGGGCACGGGGCGCGGGGCGGGCTCCCGCCCGGCACAT
GGCTGCAGCCACCTCGGCGGCACCCCGAGGCGCGCGCCAGCTCGCCCGAGGTCCGTCCGA
GGCGCCCGGCGCCGGAGCCCAAGCAGCAACTGAGGCGGGAAGCGCCCGCTCCGGGGATC
GGGATGTCCTCCTCCTTCTCCTCTTGCTAGTTTCTACTATGTTGGAACCTTGGGGACTCA
CACTGAGATCAAGAGAGTGGCAGAGGAAAAGGTCACCTTGGCCCTGCCACCATCACTGGGGC
TTCCAGAAAAAGACACTCTGGATATTGAATGGCTGCTCACCAGATAATGAAGGGAACCAAAAA
GTGGTGATCACTTACTCCAGTCGTATGTCTACAATAACTTGACTGAGGAACAGAAGGGCCG
AGTGGCCCTTTGCTTCCAATTTCTTGGCAGGAGATGCCTCCTTGAGATTGAACCTCTGAAGC
CCAGTGATGAGGGCCGGTACACCTGTAAGGTTAAGAATTCAAGGCGCTACGTTGGAGCCAT
GTCATCTTAAAAGTCTTAGTGAGACCATCCAAGCCCAAGTGTGAGTTGGAAGGAGAGCTGAC
AGAAGGAAGTGACCTGACTTTGCACTGTGAGTCATCCTCTGGCAGAGCCCATTTGTGTATT
ACTGGCAGCGAATCCGAGAGAAAAGGAGAGGATGAACGTCTGCCTCCCAAATCTAGGATT
GACTACAACCAACCTTGGACGAGTTCTGCTGCAGAATCTTACCATGTCTTACTCTGGACTGTA
CCAGTGCCACAGCAGGCCAACGAAGCTGGGAAGGAAAGCTGTGGTGCGAGTAACTGTACAGT
ATGTACAAGCATCGGCATGGTTGCAGGAGCAGTGACAGGCATAGTGGCTGGAGCCCTGCTG
ATTTTCTCTTGGTGTGGCTGCTAATCCGAAGGAAAGACAAAGAAAGATATGAGGAAGAAGA
GAGACCTAATGAAATTCGAGAAGATGCTGAAGCTCCAAAGCCCGCTCTTGTGAAACCCAGCT
CCTCTTCTCAGGCTCTCGGAGCTCAGCTCTGGTCTTCTCTCCACTCGCTCCACAGCAAT
AGTGCTCAGCGCCAGCGGACACTGTCACTGACGCGAGCACCCAGCCAGGGCTGGCCAC
CCAGGCATACAGCCTAGTGGGGCCAGAGGTGAGAGGTTCTGAACCAAAGAAAGTCCACCATG
CTAATCTGACCAAAGCAGAAACCAACCCAGCATGATCCCCAGCCAGAGCAGAGCCTTCCAA
ACGGTCTGAATTACAATGGACTTGACTCCCAAGCTTTCTAGGAGTCAGGGTCTTTGGACTC
TTCTCGTCATTGGAGCTCAAGTCACCAGCCACACAACCATGAGAGGTCATCTAAGTAGCA
GTGAGCATTGACGGGAACAGATTGAGATGAGCATTCTTCTTATACAATACCAACCAAGCAAA
AGGATGTAAGCTGATTTCATCTGTAAAAAGGCATCTTATTGTGGCTTTAGACCAGAGTAAGGG
AAAGCAGGAGTCCAAATCTATTTGTTGACCAAGACCTGTGGTGAGAAGGTTGGGGAAGGTG
AGGTGAATATACCTAAAATTTTAATGTGGGATATTTGTATCAGTGCTTTGATTACAATTT
TTCAGAGGAGACAAAATTTGTGACAAAGGATTTGTGAAGAGCTTCCATCTTCATGATGTT
ATTAGTTATTTCAGACAGTCAAGCAGAAACCCACAGCCTTATTACACCTGTCTACACCATGTAC
TGAGCTAACCACTTCTAAGAAACTCCAAAAAGGAAACATGTGTCTTCTATTCTGACTTAAC
TTCATTTGTCTAAGGTTTGGATATTAAATTCAGGGGAGTTGAAATAGTGGGAGATGGAGA
AGAGTGAATGAGTTTCTCCCACTCTATACTAATCTCACTATTGTTATTGAGCCCAAAATAAC
TATGAAAGGAGACAAAATTTGTGACAAAGGATTTGTGAAGAGCTTCCATCTTCATGATGTT
ATGAGGATGTTGACAAACATTAGAAAATATATAATGGAGCAATTTGTGGATTTCCTCCCTCAAAT
CAGATGCCTCTAAGGACTTTCTCTGATAGATATTTCTGGAAGGAGAAAAATACAACATGTCTATT
TATCAACGTCCTTAGAAAGAATTCTTCTAGAGAAAAGGGATCTAGGAATGCTGAAAGATT
CCCAACATACCATTAGTCTCTTCTTCTGAGAAAATGTGAAACCAAGATTGCAAGACTGG
GTGGACTAGAAAGGGAGATTAGATCAGTTTTCTCTTAATATGTCAAGGAAGGTAGCCGGCA
TGGTGCCAGGCACCTGTAGGAAAATCCAGCAGGTGGAGGTTGCACTGAGCCGAGATTATGCC
ATTGCACTCCAGCCTGGGTGACAGAGCGGGACTCCGCTCT

FIGURE 24

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45419

><subunit 1 of 1, 373 aa, 1 stop

><MW: 41281, pI: 8.33, NX(S/T): 3

MSLLLLLLVSYVGTLTGTHTEIKRVAEEKVTLPCHHQLGLPEKDTLDIEWLLTDNEGNQKV
VITYSSRHVYNNLTTEEQKGRVAFASNFLAGDASLQIEPLKPSDEGRYTCKVKNSGRYVWSHV
ILKVLVRPSKPKCELEGELTEGSDLTLCESSSGTEPIVYVWQRIREKEGEDERLPPKSRID
YNHPGRVLLQNLTMSYSGLYQCTAGNEAGKESCVVRVTVQYVQSIGMVAGAVTGIVAGALLI
FLLVWLLIRKDKERYEEEEERPNEIREDAEAPKARLVKPPSSSSSGSRSSRSGSSSTRSTANS
ASRSQRTLSTDAAPQPLATQAYSLVGPEVRGSEPKKVHHANLTKAETTPSMIPSQSRAFAQTV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 232-251

FIGURE 25

GTCGTTCTCTTGTCTCTCTCGCGCCAGTCCTCCTCCCTGGTTCTCCTCAGCCGCTGTGCGGAGGAGAGCACCCGGA
 GCGCGGGCTGCAGTCGCGGGCGGCTTCTCCCGCCTGGGCGGCCTCGCCGCTGGGCAAGTCTTGAGCGCCCTTAG
 AGCCTCCTCTGGCCCTTCCTCCTCTGCGCGCCGACAGTCGACATGGGGTGTGTGGAGGTAGATGGGCTCCCG
 GCCCGGGAGGCGCGGTGATGCGGCGCTGGGCAGAGCAGCCGCCGATTCCAGCTGCCCGCGCGCCCGGGG
 CCCTCGCGAGTTCGCCGTTTCAGCCATGGGGACCTCTCCGAGCAGCAGCACCGCCTCGCCTCTCTGCGCCGCTATC
 GCCCGCCGAGCCACAGCCACGATGATCGCGGCTCCCTTCTCCTGCTTGGATTTCCTTAGCACACACAGCTCAG
 CCAGAACAGAAAGGCTCGAATCTCATTTGGCACATACCGCCATGTTGACCGTGCACCGCCAGGTGCTAACTGT
 GACAAGTGTCCAGCAGGAACCTATGTCTCTGAGCATTTGACCAACACAAGCCTGCGCGTCTGACAGCAGTTGCCCT
 GTGGGACCTTTACCAAGCATGAGAATGGCATAGAGAAATGCCATGACTGTAGTCAGCCATGCCCATGGCCAATG
 ATTGAGAATTAACCTTGTGCTGCCCTTGACTGACCGAGAATGCACTTGCACCTGGCATGTTCAGATCTTAACGCT
 ACCTGTGCCCCCATACGGTGTGTCTGTGGGTTGGGGTGTGCGGAAGAAAGGACAGAGATGAGGATGTGCGG
 TGTAAAGCAGTGTCTCGGGGTACCTTCTCAGATGTGCCTTCTAGTGTGATGAAATGCAAGCATACACAGACTGT
 CTGAGTCAGAACCTGGTGGTGATCAAGCCGGGACCAAGGACAGACAACCTCTGTGGCACACTCCCGTCTCTC
 TCCAGCTCCACCTCACCTTCCCTTGGCACAGCCATCTTTCCACGCGCTGAGCACATGGAACCCATGAAGTCCCT
 TCCTCCACTTATGTTCCCAAGGCATGAACCTCAACAGAATCCAATCTTCTGCCTCTGTAGACCAAGGTACTG
 AGTAGCATCCAGGAAGGSCAGTCCCTGACCAACACAGCTCAGCAAGGGGGAGGAGAACGTGAACAGACCCCTC
 CCAAACTTCAGGTAGTCAACACACAGCAAGGCCCCACACAGACACATCTGAAGCTGCTGCCGTCCATGGAG
 CCACTCGGGGCGAGGAAGTCCAGACAGCCCATCAAGGGCCCCAAGAGGGGACATCTGACAGCAAGCTACACAAG
 CATTTTGACATCAATGAGCATTTGCCCTGGATGATTGTGCTTTTCTGCTGCTGTGTGCTGTGTGTGATTGTGGT
 TGCAGTATCCGGAAGGCTCGAGGACTCTGAAAAAGGGGCCCGCGAGGATCCGATGCCATTTGTGAAAAAGGCA
 GGGCTGAAGAAATGATGACTCAACACCGAGACCGGGAGAAATGGATCTACTACTGCAATGGCCATGGTATCGAT
 ATCTTGAAGCTTTCAGACCCCAAGTGGGAAGCCAGTGGGAAGATATCTATCAAGTTCTTTGCAATGGCAGTGAG
 AGGGAGGTGTGCTGCTTTCTCAATGGGTACACAGCGGACCACGAGCGGGGCTACGCACTCTCGCAGCATGGAGC
 ATCCGGGGGCCCGAGGCCAGCTCGCCAGCTAATTAGCGCCCTGCGCGAGCACCGGAGAAACGATGTGTGGAG
 AAGATTCTGGGGTGTGGAAGACACACCCAGCTGGAACATGCAAACTAGCTCTCCGATGAGCCCGAGCCG
 CTTAGCCCGAGCCCATCCCGAGCCCAACGCGAAAATTGAGAAATTCGCTCTCTGACCGGTGGAGCCTTCCCA
 CAGGACAAGAACAGGGCTTCTCGTGGATGAGTCGGAGCCCTTCTCCGCTGTGACTCTACATCAGGGGCTCC
 TCCGCGCTGAGCAGGAACGGTTCTTTATTACCAAGAAAAAGAGGACACAGTGTGTGGCGAGGTACGCTGGAC
 CCCTGTGACTTGCAGCCTATCTTTGATGACATGCTCCACTTCTAAATCTCGAGGAGCTGCGGGTGATTGAAGAG
 ATTCCCGAGGCTGAGGACAACTAGACCGGCTATTGCAAAATATTGGAGTCAAGAGCCAGGAAGCCAGCCAGACC
 CTCCTGGACTCTGTTTATAGCCATCTCCGTGACCTGCTGTAGAACATAGGGATACGCAATCTGGAATTAATCA
 ATTTAGTGGCAGGGTGTTTTTAAATTTCTTCTGTTTTGATTTTTGTTGTTTGGGGTGTGTGTGTGTGTTTTGT
 GT
 TCTCTCTTTTTTTTTTAAATAAATCTCTGCGGAAGTTGGTTTTATAAGCCTTTGCGAGGTGTAATCTGTGTGAA
 ATACCCACCACTAAAGTTTTTAAAGTCCATATTTCTCCATTTTGGCCTCTTATGATTTTCAAGATTATTCTGT
 TGCACTTTAAATTTTACTTAACTTACCATAAATGCAGTGTGACTTTCCCAACACACTGGATTGTGAGGCTCTTAAC
 TTCTTAAAGTATAATGGCATCTTGTGAATCTATAAGCAGTCTTATGTCTCTTAAACATTACACCTACTTTTT
 AAAAAAATAATTATTAATCTATTTTATTATTGTTGTCCTTATAAAATTTCTTAAAGATTAAAGAAATTAAGAA
 CCCCATTGAGTTACTGTAAATGCMAATCAACTTTGAGTTATCTTTAAATATGCTTTGATATGTTCATATTCAATG
 CTGAAACTTGAACCACTATGCTGATGTGTATGGTTTTCACCTGGACACCGGTGAGAATGCTGATATTCTGTATC
 TCTTCTTATGCTAATGCTCTGGGCTGGAGAAATGAAATCCTCAAGCATCAGGATTTGCTATTTAAGTGGCTT
 GACAACCTGGCCACCAAGAACTTGAACCTTCACTTTTAGGATTTGAGCTGTCTGGAACACATCTGCTGCACCTT
 GGAAGCTCAAAATCAAGTGCAGTGGCGCCCTTCCATAGAGAATTTGCCGACCTTTGCTTTAAAGAGTGTCTG
 TTTTATATACATAATCAATAATAGTCCAATCTGCTCTCAAGGCTTGGTCTGTGGGATCCCTTCAAAAT
 ACTTTAAATTAAGATGGTGCMACTGTAAAGACCCCTTGTCTGATATTTTGCACCAATGTCTCCATTTACAAATG
 TACCTTCAATGCTGACTGTGCCAGGTCCAAATGCAAGGGTGGCGTGGACTCCCTTTGTGTGGTGGGGTTGTGG
 GTAGTGGTAAGGACGATACAGAAAAATGCCCTCAAGTGATCAATTTATTAATAAACATTAGGTGTTGTGTA
 AAAAAAAA

FIGURE 26

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52594
><subunit 1 of 1, 655 aa, 1 stop
><MW: 71845, pI: 8.22, NX(S/T): 8
MGTSPPSSSTALASCSRIARRATATMIAGSLLLLGFLSTTTTAQPEQKASNLIQTYRHVDRATG
QVLTCDKCFAGTYVSEHCTNTSLRVCSSCPVGTFTTRHENGIEKCHDCSQPCPWPMIEKLPCA
ALTDRECTCPPGMFQSNATCAPHTVCPVGWGVRRKGTETEDVRCKQCARGTFSDVPSSVMKC
KAYTDCLSQNLVVIKPGTKETDNVCGTLPSFSSSTSPSPGTAI FPRPEHMETHEVPSSTYVP
KGMNSTESNSSASVREKVLSSIQEGTVPDNTSSARGKEDVNKTLPNLQVVNHQQGPHHRHIL
KLLPSMEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLPWMIVLFLLLVLVVIVVCSIRK
SSRTLKKGPRQDPSAIVEKAGLKKSMPTPTQNREKWIYYCNGHGIDILKLVAQVGSQWKDIY
QFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISALRQHRNDVVEKIRG
LMEDTTQLETDKLALPMSPSPLSPSPISPNAKLENSALLTVEPSQDKNKGFFVDESEPLL
RCDSTSSGSSALSRRNGSFITKEKKD TVLRQVRLDPCDLQPIFDDMLHFLNPBELRVIEEIPQ
AEDKLDRLFEIIGVKSQEASQTLLDSVYSHLPDLL
```

Signal sequence:

amino acids 1-41

Transmembrane domain:

amino acids 350-370

FIGURE 27

ATGGGAAGCCAGTAACACTGTGGCCTACTATCTCTTCGTTGGTGCCATCTACATTTTTTGGGA
CTCGGGAAATATGAGGTAGAGGTGGAGGCCGAGCCGGATGTCAAGGTCCTGAAATAGTCAC
CATGGGGGAAAAATGATCCGCTGCTGTTTGAAGCCCCCTTCTCATTCGGATCGCTTTTGGCC
TTGATGATTTGAAAAAAGTCCTGTTGCAACAGATGCAGATGCTGTTGCTGCAAGATCCTG
TCACTGCTGCCATTGAAGTTTTTTCCAATCATCGTCATTGGGATCATTGCAATGATATTAGC
ACTGGCCATTGGTCTGGGCATCCACTTCGACTGCTCAGGGGAAGTACAGATGTGCTCATCCT
TTAAGTGTATCGAGCTGATAGCTCGATGTGACGGAGTCTCGGATTGCAAGACGGGGAGGAC
GAGTACCGCTGTGTCCGGGTGGGTGAGTGCAGTGCCTGCTCAGGTTGTTACAGCTGCTTTC
GTGGAAGACCAATGTGCTCCGATGACTGGAAGGGTCACTACGCAAAATGTTGCTGTGCCCCAAC
TGGGTTTCCCAAGCTATGTGAGTTAGATAACCTCAGAGTGAGCTCGCTGGAGGGGCAAGTTC
CGGAGGAGTTTGTGTCCATCGATCACCTCTTGGCAGATGACAAGGTGACTGCATTACACCA
CTCAGTATATGTGAGGGAGGGATGTGCCTCTGGCCACGTGGTTTACCTTCAGTGACAGCCT
GTGGTCATAGAAGGGGCTACAGCTCACGCATCGTGGGTGGAACATGTCTTGTCTCTCGCAG
TGGCCCTGGCAGGCCAGCCTTCAGTTCAGGGGTACACCTGTGTCGGGGGCTCTGTATCAC
GCCCCGTGTGGATCATCACTGCTGCACACTGTGTTTATGACTTGTACCTCCCCAAGTCATGGA
CCATCCAGGTGGGTCTAGTTTCCCTGTTGGACAATCCAGCCCCATCCCCTTGGTGGAAGAAG
ATTGTCTACCACAGCAAGTACAAGCCAAAGAGGCTGGGCAATGACATCGCCCTTATGAAGCT
GGCCGGGCCACTCACGTTCAATGAAATGATCCAGCCTGTGTGCTGCCCACTCTGAAGAGA
ACTTCCCGCATGGAAGAAGTGTCTGGACGTGAGGATGGGGGGCCACAGAGGATGGAGGTGAC
CCCTCCCCTGTCTGAACCAAGCGCGGCCGTCCCTTTGATTTCACCAAGATCTGCAACCCAG
GGACGTGTACGGTGGCATCATCTCCCCCTCCATGCTCTGCGCGGGCTACCTGACGGGTGGCG
TGGACAGCTGCCAGGGGGACAGCGGGGGGCCCTGTTGTCTCAAGAGAGGAGGCTGTGGAAG
TTAGTGGGAGCGACAGCTTTGGCATCGGCTGCGCAGAGGTGAACAAGCCTGGGGTGTACAC
CCGTGTCACTCTCTCTGGAAGTGGATCCACGAGCAGATGGAGAGAGACCTAAAAACCTGAA
GAGGAAGGGGACAAGTAGCCACCTGAGTTCTGAGGTGATGAAGACAGCCGATCCTCCCT
GGACTCCCGTGTAGGAACCTGCACAGAGCAGACACCCTTGGAGCTCTGAGTTCCGGCACC
GTAGCAGGCCCGAAAGAGGACACCCTTCCATCTGATTCCAGCACACCCTTCAAGCTGCTTTTT
GTTTTTTGTTTTTTTGGAGTGGAGTCTGCTCTGTTTGCCAGGCTGGAGTGCAGTGGCGAAA
TCCCTGCTCACTGCAGCCTCCGCTTCCCTGGTTCAAGCGATTCTCTTGCTCAGCTTCCCCA
GTAGCTGGGACCAAGGTGCCGCCACCAACCACTAATTTTTGTATTTTTAGTAGAGAC
AGGTTTACCATGTTGGCCAGGCTGCTCTCAAAACCCTGACCTCAAATGATGTGCTGCTT
CAGCTCCACAGTGTGAGGATTACAGGCATGGGCCACCAAGCCTAGCCTCAGCTCCTTTC
TGATCTTCACTAAGAAACAAAGAAGCAGCAACTTGCAAGGGCGGCTTCCCACTGGTCCAT
CTGGTTTTCTCTCCAGGGTTCTGCAAAATCTGTACGAGATAAGCAGTTATGTGACCTCAG
TGCAAGGCCACCAACAGCCTCTCAGAAAAGACGACCCAGCCAGAGTGCAGAACTGCAGTC
ACTGCACGTTTTTCATCTCTAGGGACCAGAACCAACCCACCTTCTACTTCCAAGACTTAT
TTTCAATGTGGGGAGGTTAATCTAGGAATGACTCGTTTAAAGGCTATTTTCATGATTTCTT
TGTAGCATTTGGTGCTTGAAGTATTTATGTCCTTGTGATTCCAAATATGTTTCTTCCCT
CATGTCTGGCGTGTCTGCTGGAAGTGGTGACGTGAATCAAATCATCCACTGAAA

FIGURE 28

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45234

><subunit 1 of 1, 453 aa, 1 stop

><MW: 49334, pI: 6.32, NX(S/T): 1

MGENDPPAVEAPFSFRSLFGLDDLKISPVAPDADAVAAQILSLLPLKFFPIIVIGIIALILA
LAIGLGIHFDCSGKYRCRSSFKCIELIARCDGVSDCKDGEDEYRCVVRVGGQNAVLQVPTAAS
WKTMCSDDWKGHYANVACAQLGFPSYVSSDNLRVSSLEGQFREEFVSIHLLPDDKVLTALHH
SVYVREGCAGSHVVTLQCTACGHRRGYSSRIVGGMNLSLQWPWQASLQFQGYHLCGGSVIT
PLWIIITAAHCVDYDLYLPKSWTIQVGLVSLLDNPAPSHLVEKIVYHSKYKPKRLGNDIALMKL
AGPLTFNEMIQPVCLPNSEENFPDGGKVCWTSGWGATEDGGDASPVLNHAAVPLISNKICNHR
DVGGIISPSMLCAGYLTGGVDSCQGDSGGPLVCQERRLWKLVGATSGFIGCAEVNKPVGVT
RVTSFLDWIHEQMERDLKT

Signal Peptide:

amino acids 1-20

Transmembrane domain:

amino acids 240-284

FIGURE 29

CCCACGCGTCGTCCTAGTCCCOCGGGCCAACTCGGACAGTTTGCTCATTATTATGCAACGGTCAAGGCTGGCTTGT
 GCCAGAAACGGCGCGCGCGCGCGCACGACGCGCACACACGGGGGGAACCTTTTTTAAATAAGAAAGGCTAGAAAG
 GCTCAGCGCGCGCGCGCGCGCTGCGCGAGGGCTCGGAGCTGACTCGCGGAGCAGGAAATCCCTCCGCTCGCGA
 CGCCCGGCCCGCGCTCGCGCGCCGCTGGGATGGTGACGCGCTCGCGCGCGGGCCGAGAGCTGCTGCACTGAAG
 GC CGCGGACGATGCGACGCGCGCGCTGCCGCTGTCGCCCGCGCCCTCTGCTCGCCCTGGCCGCTGCTCT
 GCTCGCGCCCTGCGAGGCGCGAGGGGTGAGCTTATGGAACCAAGGAAGAGCTGATGAAGTTGTCACTGCTCTGT
 TCGGAGTGGGACCTCTGGATCCAGTGAAGAGCTTCGACTCCAAGAATCATCCAGAAGTGTCTGAATATTCGACT
 ACAACGGGAAAGCAAGAACTGATCATAAATCTGGAAGAAATGAAGTCTCATTTGCCAGCAGTTTACCGGAAAC
 CCACTATCTGCAAGACGCTACTGATGCTCCCTCGCTCGAAATACACGGGTCACTGTTACTACCATCGGACATGT
 ACGGGATCTTCTGATTCAGCAGTCACTCTGAGCAGCTGTTCTGGTCTCAGGGGCTATTGTTGTTTGAATAATGA
 AAGCTATGTTTACAGCAATGAAAAGTGAACCAACAGATACAACTCTTCCAGCAGGAAGCTGAAAAGCGT
 CCGGGATCATGTGGATCACATCAACACACCAAACTCGCTGCAAGAAATGTGTTTCCACCAACCTCTCAGAC
 ATGGGCAAGAGGCTATAAAGAGAGACCTCAAGGCAACTAAGTATGTGGAGCTGGTGATGCTGGCAGACAAACCG
 AGAGTTTTCAGAGCAAGGAAAAGATCTGGAAGAAAGTTAAGCAGCGATTAAAGAGATGTCTTAATCACTGTGACAA
 GTTTTACAGACCACTGAACATTTCGATCGTGTGGTAGGCGTGGAGTGTGGATGACATGGACAAATGCTCTGT
 AAGTCAGGACCCATTCAACGACCTCCATGAATTTCTGGACTGGAGGAAGATGAAGCTTCTACCTCGCAATCCCA
 TGACAATGCGCAGCTTGTCACTGGGGTTTATTTCAAGGGACACCATCGGCATGGCCCAATCATGAGCATGTG
 CACGGCAGACCACTGTGGGGGAATTTGTCATGGACCATTCAGACAAATCCCTTGGTGAGCCGTGACCTGGGCACA
 TGAGCTGGGCCAAATTTTCGGGATGAATCATGACACACTGGACAGGGCTGTAGCTGTCAAATGGCGGTTGAGAA
 AGGAGGCTGCATCATGAACGCTTCCACGGGTACCCATTCCCATGGTGTTCAGCAGTTTGACGAGGAAGGACTT
 GGAGACAGGCTGAGAAAGGAATGGGGGTGTGCTGTTTAACTGCCGGAAGTCAGGGAGTCTTTCGGGGGCCA
 GAAGTGTGGGAACAGATTTTGTGAAGAAGGAGAGGAGTGTGACTGTGGGGAGCCAGAGGAATGTATGAATCGCTG
 CTGCAATGCCACACCTGTACCTGAAGCGGACGCTGTGTGCGACATGGGCTGTGCTGTGAAGACTGCCAGCT
 GAAGCTCGAGGAACAGCGTGCAGGCACTCCAGCAACTCCTGTGACCTCCCAAGATTTCTGCAAGGGGGCCAGCC
 TCACTGCCAGCCAAATGTGTACCTGCAGATGGGCACTCATGTCAAGATGTGGAAGGCTACTGTCTAATGGCAT
 CTGCCAGACTCAGAGCAGCAGTGTGTCACTCTGGGGACAGGTGCTAAACCTGCCCTGGGATCTGCTTTGA
 GAGAGTCAATTCGACAGTGTCTTATGGCAACTGTGGCAAGTCTGGAAGAGTTCCTTTGCCAAATGCGAGAT
 GAGAGATGCTAAATGTGGAAGAAATCCAGTGTCAAGGAGGTGCCAGCGGCCAGTCAATGGTACCAATGCCGTTTC
 CATAGAAACAAACATCCCTCTGACAGCAGGAGGCGGAGTTCTGTGCGGGGGACCCACGCTGATCTTGGCGATGA
 CATGCGGACCCAGGGCTTGTCTGTGAGGCAAAAGTGTGAGATGGAAGAAATCTGCGTGAATCGTCAATGTGA
 AAATATTAGTGTCTTTGGGTTTCAGCAGTGTGCAATGCACTGCCAGCAGAGGGGTGTGCAACCAACAGGAAGAA
 CTGCCACTGCGAGGCCCATCTGGGCACTCCCTTCTGTGACAAAGTTTGGCTTTGGAGGAAGCAGACAGCGGCC
 CATCGGCAAGCAGAGCAGGCGAGGAGCTGACAGTCCAACAGGAGCGCGGCGAGGGCCAGGAGCCCGTGGG
 ATCGCAGGAGCATGCTCTACTGCGCTCACTGACACTCATCTGAGGCCCTCCCATGACATGGAGCCGTGACCAATG
 CTGCTCGAGAGGAGGTCAAGCTGCCCAAGGCCCTCTGTGACTGGCAGCATGACTCTGTGGCTTTGCCATCGTT
 TCCACTCAACAGACAGACAGTCTGCGGCTCAGGAGGGGAAGTCCAGCTCTCCAGGCAGTCTGCAAGAA
 CAGTGAAGGAGGAGGACGCACTCTCTGGTGTGAGCTTCTGCTAAACATGGACATGCTTCACTGCTGCTCTGAG
 AGATGAGGAGTTTACACTCTGCGAGGCCCGACGCTCGACGAGGAAGGAATCAAGATCTGCTGCTGCTGCTGCTT
 ACTGAGCCCTCAGACAGTGGGGGAGAGCAAGGGTGTGGCCGAGTTCCTCTTCCCACTGACACCTCAGCCT
 TGGCAGCCCTGACAGTGGTCTCTGAGCTCAACTTAATGCTCTGATATGGCTTTTAGCATTTATATATAGAAAT
 AGCAGGGTTTATGTTTAAATTTATCAGAGACCTGCCACCAATCCATCTTCCATCAAGCAAACTGAATGCAAA
 TGAACCAACTGGAGGAAGATGAGTAGGAAGGGCGGTGAATCTGGGCTCTTCTGTGGACATGCTGACAGCAG
 AGTACTCAGGTTTGGGGTTTGCAGAAAGCCAGGGAAACCCACAGAGTCAACCAACCTTCAATTAACAGTAAAGAA
 TGTTTAAAGAGTGAACCAATGTAAGAGCCCTAACTCCATCCCCGTTGGCCATTACTGCATAAATAGAGTGCATT
 GAAAT

FIGURE 30

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49624

><subunit 1 of 1, 735 aa, 1 stop

><MW: 80177, pI: 7.08, NX(S/T): 5

MAARPLPVSPARALLLALAGALLAPCEARGVSLWNQGRADEVVVSASVRSGLWIIPVKSFDSK
NHPEVLNIRLQRESKELIINLERNEGLIASSFTETHYLQDGTDVSLARNYTGHCYYHGHVRG
YSDSAVSLSTCSGLRGLIVFENESYVLEPMKSATNRYKLFPAKKLKSVRGSCGSHHNTPNLA
AKNVFPPPSQTWARRHKRETLKATKYVELVIVADNREFQRQGDLEKVKQRLIEIANHVDFK
YRPLNIRIVLVGVEVWMDMKCSVSQDPFTSLHEFLDWRKMKLLPRKSHDNAQLVSGVYFQG
TTIGMAPIMSMCTADQSGGIIVMDHSDNPLGAAVTLAHELGHNFGMNHDTLDRGCSCQMAVEK
GGCIMNASTGYPPFMVFSSCSRKDLETSLEKMGVCLFNLPEVRESFGGQKCGNRFVEEGEE
CDCGEPEECMNRCNATTCTLKPDCAVCAHGLCCEDCQLKPAGTACRDSNSNCDLPEFCTGAS
PHCPANVYLHDGHSQDQVDGYCYNGICQTHEQQCVTLWGPAGAPAGICFERVNSAGDPYGN
CGKVSXSFAKCEMRDAKCGKIQCGGASRPVIGTNAVSIETNIPLQQGGRILCRGTHVYLG
DDMPDPGLVLGATKCADGKICLNRCQCNISVFGVHECAMQCHGRGVCNNRKNCHCEAHWAPP
FCDKFGFGGSTDSGPIRQAEARQEAESNRERGGQGEVPGSQEHASTASLTLI

Signal peptide:

amino acids 1-28

FIGURE 31

Figure 1

FIGURE 32

CATCCTGCAACATGGTGAAACCACGCCTGGCTAATTTTGTGTATTTTGGTAGAGATGGGA
TTTCACCGTGTTAGCCAGGATTGTCTCAATCTGACCTCATGATCTGCCCGCTCGGCCTCCC
AAAGTGCTGGGATTACAGGCGAGTGCAACCACACCCGGCCACAACTTTTAAAGAAGTTAAT
GAAACCATACTTTTACATTTTAAATGACAGGAAAATGCTCACAATAATTGTTAACCCAAA
TTCTGGATACAAAAGTACAATCTTTACTGTGTAAATACATGTATATGTACTATATGAAAATA
TACCAAATATCAATAATACTTATCTCTGGGTAAAACCTCTTCTCATACCCTGTGCTAACAA
CTTTTAACAAAAATTTGCATCACTTTTAAAGAAATCAAGAAAAATTTCTGAAGGTCAATGGG
ACAGAAAAAAAACCAAGGGAAAAATCACGCCACTTGGGAAAAAAGATTGCAAAATCTGCCT
TTTTATAGATTTGTAATTAATAAGGTCCAGGCTTTCTAAGCAACTTAAATGTTTGTTCGA
AACAAAGTACTTGTCTGGATGTAGGAGGAAAGGAGTGATGTCACTGCCATTATGATGCCCC
TTGAATATAAGACCCCTACTTGTCTATCTCCCCTGCACCAGCCAGGAGCCACCCATCCTCCAGC
ACACTGAGCAGCAAGCTGGACACACGGCACACTGATCCAAATGGGTAAGGGATGGTGGCGA
TGCTCATTCTGGGTCTGCTACTTCTGGCGCTGCTCCTACCCGTGCAGGTTTCTTCATTTGTT
CCTTTAACAGTATGCCGGAAGCTACTGCAGCCGAAACCACAAAGCCCTCCAACAGTGCCCT
ACAGCCTACAGCCGGTCTCCTTGTGGTCTTGCTTGCCCTTCTACATCTCTACCATTAAAGAGG
CAGGTCAAGAAACAGCTACAGTTCTCCAACCCATACACTAAAACCGAATCCAAATGGTGCCCT
AGAAGTTCAATGTGGCAAGGAAAAAAACCAGGTCTTCATCAAATCTACTAATTTCACTCCTT
ATTAACAGAGAAACGCTTGAGAGTCTCAAACCTGGACTGGTTTAAAGAGCATCTGAAGGATT
GACTAGATGATAAATGCCTGTACTCCAGTACTTTGGGAGGCCTAGGCCGGCGGATCACCTG
AGGTGAGGAGTTTGAGACTAACCTGGCCAAAATGGTGAAACCCCATCTGTACTAAAAATACA
AATATTGACTGGGCGTGGTGGTGAGTGCCGTGTGATCCAGCTACTCAGGTGGCTGAAGCAGG
ACAATCACTTGAACCTCAGGAGGCAGAGGTTGCAGTGAGCTGAGATCGCGTACTGCACCTCA
GCCTAGCCTGGGCAACAGAGTGAGACTTCGTCTCAAAAAAAAAAAGCCAAGTGACGTGGCT
CACGCTGTAATCCCGGCACCTTTGGGAGGCCGAGGTGGGCGGATCACGAGGTCAGGAGATCA
AGACCATCCTGGCTAATACAGTGAACCCCTGTCTCTACTAAAAATAAAAAAATTAGCCGGG
GATGTTGGCAGGCACCTGGAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAAATAGCGTGAA
CTCAGGAGGCGGAGCTTGCAGTGAGCCGAGATTGCGCTACTGCATCCAGCCTGGGCGACAG
CGCGAGACTCCGTCTCAAAAAAAAAAAAAAAAAAAAAA

FIGURE 33

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48309

><subunit 1 of 1, 67 aa, 1 stop

><MW: 6981, pI: 7.47, NX(S/T): 0

MGKGMVAMLILGLLLLALLLPVQVSSFVPLTSMPEATAAETTKPSNSALQPTAGLLVLLAL
LHLYH

Signal peptide:

amino acids 15-27

100
90
80
70
60
50
40
30
20
10
0
100
90
80
70
60
50
40
30
20
10
0

FIGURE 35

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA46776

><subunit 1 of 1, 432 aa, 1 stop

><MW: 47629, pI: 5.90, NX(S/T): 0

MPARPGRLLPLLARPAALTALLLLLLGHGGGGRWGARAQEAAAAAADGPPAADGEDGQDPHS
KHLYTADMFTHGIQSAAHFVMFFAPWCGHCQRLQPTWNDLGDKYNSMEDAKVYVAKVDCTAH
SDVCSAQGVRGYPTLKLFKPGQEAVKYQGPRDFQTLNWMQLTLNEEPVTPEPEVEPPSAPE
LKQGLYELASNFELHVAQGDHFIKFFAPWCGHCKALAPTWEQLALGLEHSETVKIGKVDCT
QHVELCSGNQVRGYPTLLWFRDGGKVDQYKGKRDLESLEYVESQLQRTETGATETVTPSEA
PVLAAEPEADKGTVLALTENNFDDTIAEGITFIKFYAPWCGHCKTLAPTWEELSKKEFPGLA
GVKIAEVDCTAERNICSKYSVRGYPTLLLFRGGKKVSEHSGGRDLDSLHRFVLSQAKDEL

Signal sequence:

amino acids 1-32

FIGURE 36

CTTTTCTGAGGAACACAGCAATGAATGGCTTTGCATCCTTGCTTCGAAGAAACCAATTTAT
CCTCCTGGTACTATTTCTTTTGCAAATTCAGAGTCTGGGTCTGGATATTGATAGCCGTCCTA
CCGCTGAAGTCTGTGCCACACACAATTTACCAGGACCCAAAGGAGATGATGGTGAAAAA
GGAGATCCAGGAGAAGAGGGAAGCATGGCAAAGTGGGACGCATGGGGCCGAAAGGAATTAA
AGGAGAACTGGGTGATATGGGAGATCAGGGCAATATTGGCAAGACTGGGCCCATTTGGGAAGA
AGGGTGACAAAGGGGAAAAAGGTTTGCTTGGAATACCTGGAGAAAAAGGCAAAGCAGGTACT
GTCTGTGATTGTGGAAGATACCGGAAATTTGTTGGACAACCTGGATATTAGTATTGCTCGGCT
CAAGACATCTATGAAGTTTGTCAAGAATGTGATAGCAGGGATTAGGGAACTGAAGAGAAAT
TCTACTACATCGTG CAGGAAGAGAAGAACTACAGGGAATCCCTAACCCTGCAGGATTCGG
GGTGGAATGCTAGCCATGCCAAGGATGAAGCTGCCAACACACTCATCGCTGACTATGTTGC
CAAGAGTGGCTTCTTTCGGGTGTTCAATTGGCGTGAATGACCTTGAAAGGGAGGGACAGTACA
TGTCCACAGACAACACTCCAATG CAGAACTATAGCAACTGGAATGAGGGGAACCCAGCGAC
CCCTATGGTCATGAGGACTGTGTGGAGATGCTGAGCTCTGGCAGATGGAATGACACAGAGTG
CCATCTTACCATGTACTTTGTCTGTGAGTTCATCAAGAAGAAAAAGTAACTTCCCTCATCCT
ACGTATTTGCTATTTTCTGTGACCGTCATTACAGTTATTGTTATCCATCCTTTTTTTCCTG
ATTGTACTACATTTGATCTGAGTCAACATAGCTAGAAAATGCTAAACTGAGGTATGGAGCCT
CCATCATCAAAAAAAAAAAAAAAAAA

FIGURE 37

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50980

><subunit 1 of 1, 277 aa, 1 stop

><MW: 30645, pI: 7.47, NX(S/T): 2

MNGFASLLRRNQFILLVLFLLQIQSLGLDIDSRPTAEVCATHTISP GPKGDDGEKGD PGEEG
KHGKVG RMGPKGIKGELGMDGQGNIGKTGP IGGKGD KGEKGLLGIPGEK GKAGTVCD CGRY
RK FVGQLDISIARLKTSMKFVK NVIAGIRETEEFY I VQEKNYRESLTHCRIRGGMLAMP
KDEAANTLIADYVAKSGFFRVFIGVNDL EREGQYMSTDNTPLQNYSNWNEGEPSDPYGHEDC
VEMLSSGRWNDTECHLTMYFVCEFIKKKK

Signal peptide:

amino acids 1-25

GGTTCATCGATTGCAATTCGGCCCACTGGCGGATCCTCTAGAGATCCCTCGACCTCGAC
CCACGCGTCCTCGCTGCTCTCCCGCCGTTGGAGTGTGGGGGCTGGGTGGGAATCGCGGTGT
GCCAGCGCACGCGCGCTCCCTGGAAGGAGAAGTCTCAGCTAGAACGACGCGCCTTAGTTT
CGGAAGGGAGGATCAGGAGTGTTCGCGAGCGGCTGGAACCCAGACGCGTCCGATAGAGGAAGC
GGGCTCCAGTGGCTGCCCTCTGCTGCTGCCCTGCTGCTGTTGCTACCGCTGCTGCTGCTGA
AGTCAACCTCTGGCGCGAGTTTGCCTGGCTCTCCGCGCACTTGGCCCTTTCGGTGGCAGCT
CTGTGCTCGAAAAGGGCTCTTCAGAGCTCGCGCTTCGCGCGGCTCGCGCGACCCGCGAAGG
TCCGAGAGGGGGGCTCGACGCTGGCTGGCGCTCGCGAACCCTGGCCACGACGCGCGCGCG
ACACCTTTTCTCATTACGGCTCGCGGCGCTTTAGCTACTCAGAGGCGGAGCGCGAGAGTAA
AGGGCTGGACGCGGCTTCTCAGTGCCTTAGGCTGGGATCTGGGACCCGACGCGCGCGACAG
CGCGGAGGGGAGCGCTGGAGAAGGCGACGGCGAGCGCGGAGCCGAGATCGACGCGCGCG
GAAGCGCGCGCGGAGTTTCGCGGAGGGGACGCTGCCGCGCAGAGGTTGAGAGAGTCCGCCCCCT
CTGTCACCTGGAGCAACTGTGGCGCTGCTCTCTCCCGCTGGCCGAGAGTTTCTGTGGCTGTG
GTTGGGGCTGGCCAAAGGCGCGCTGCGCATGCTTTTGTGCCACCGGCTTTCGCGCGGGGCG
CCCTGCTGCAGCTGCTCCCGAGCTCGCGCGCGCGCGCTGTGTGTCGGCGCCAGAGTTTCTG
GAGTCCCTGGAGCGGCACTGCGCGCTCGAGACGATGGGCTCCACCTGTGGCTGCAAG
CCCAGAACCCCACTCTGTGGAATTAGCGATTGTGCTGGCTGAAGTGTCCGCTGAAGTGGATG
GCCGAGTGGCCAGGATACCTCTCTCCCGCCAGAGCATAAACAGACAGTGCCTGTACATCTTC
ACCTCTGGCACCAAGGCGCTCCCAAGGCTGCTCGGATCAGTCATCTGAAGATCTCGTAAGT
CCAGGGCTTCTATCAGCTGTGTGGTGTCCACCAAGGAAGATGTATCTACCTGCCCTCCCA
CTACCCACATGTCCGGTTCCCTGCTGGGCATCTGGGCTGCAATGGGCAATGGGCGCACATG
GTGCTGAATCCAAGTTTCTCGGCTGGTCAGTTCTGGGAAGATTGCCAGCAGCAGGGTGAC
GGTGTTCAGTACATATGGGGAGCTGTGCCGATACCTTTCAACACGCCCCAGCAGAGGCAG
AACGTGGCCATAAGTTCGGCTGGCATGGTGGAGCGGGCTGCGCCAGATACCTGGGAGCGT
TTTGTGCGCGCGCTTCGGGCGCTGCGAGTGTGAGACATGGACATGACAGAGGGCAACGT
GGCCACCATCAACTACACAGGACAGCGGGGCGCTGTGGGCGTGCTTCCTGGCTTTACAAG
ATATCTTCCCTTCTCTCTGATTGCTATGATGTACACAGAGAGAGCCAACTCGGGAACCC
CAGGGGCACTGTATGGCCATCTCAGGTGAGCGAGGCTGCTGGTGGCCCGGTGATAGCCA
CGAGTCCCAATCTCTGGCTATGCTGGCGGCGAGAGCTGCCACGGGGAAGTTGCTAAAGC
ATGTTCTTCGGCTTGGGATGTTTCTTCAACACTGGGACCTGTGTTGTGCGATGACCAA
GGTTTTCTCCGCTTCCATGATCGTATCGTAGAGACACCTTCAGGTGGAGGGGAGAGTGTGC
CACAAACGAGGTTGGCAGAGGTTCTCAGGCGCTAGATTTTCTCAGAGGTTGAACGTCTATG
GAGTCACTGTGCCAGGCACTAAGGACAGGCTGGAATGGCAGCCCTAGTTTCTGCTGCTCCCC
CACGCTTTGGACCTATGCACTCTACACGCTGACGCTGTCTGAGAACTTGGCACTTATGCGCG
GCCCGGATTCTCAGGCTCCAGAGTCTTTGGCCACCACAGAGACCTTCAAACAGCAGAAAG
TTCGGATGGCAAATGAGGGCTTCGACCCAGCACCTGTCTGACCACCTGTACGTTCTGGAC
CAGCTGTAGGTGCCATCTCGCTCCCTCAACAATGCCGCTGACGAGCGGCTTCGCGCAGAGAA
CCTTCGAATCTGAAGACTTCCACACTCAGGCACTGAGAGACCTGAGAGAGAACTCTGTGGGTGGGG
CGTGTGTCAGGTGTACCTGGGCTGTCAAGGATCTTTTCTATACCAGAGTCGGTCACTATTTT
GTAATAAATGTGGCTGGAGCTGATCCAGCTGTCTCTGACCTAAAAAAGGAAAAAAGGAAAA
AAAAAAGGAGGCGCGGCACTTAGAGTGCACCTGAGTAGGGAATAACAGGGTAAATAAGC
TTGGCGGCATTCGCGCAACTTGTTTATTGCGAG

FIGURE 39

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50913

><subunit 1 of 1, 730 aa, 1 stop

><MW: 78644, pI: 7.65, NX(S/T): 2

MGVCQRTAPWKEKSQLERAALGFRKGGSGMFASGWNQTVPIEEAGSMAALLLPLLLLLP
LLLLKLHLWPQLRWLPADLAFAVRALCCKRALRARALAAAAADPEGPEGCSLAWRLAELAQQ
RAAHTFLIHGSRFFSYSEAERESNRAARAFRLRALGWDWGPDDGDSGEGSAGEGERAAPGAGD
AAAGSGAEFAGGDGAARGGGAAAPLSPGATVALLLPAGPEFLWLWFLAKAGLRTAFVPTAL
RRGPLLHCLRSCGARALVLAPEFLESLEPDLPALRAMGLHLWAAGPGTHPAGISDLLAEVSA
EVDGVPVGYLSSPQSITDTCLYIFTSGTTGLPKAARISHLKILQCQGFYQLCGVHQEDVIYL
ALPLYHMSGSLGIVGCMGIGATVVVKSKFSAGQFWEDCQHRVTVFQYIGELCRYLVNQPP
SKAERGHKVR LAVGSGLRPDTWERFVRRFGPLQVLETYGLTEGNVATINYTGQRGAVGRASW
LYKHIFPFSLIRYDVTTEPIRDPQGHCMATSPGEPGLLVAPVSQQSPFLGYAGGPELAQ GK
LLKDVFVRPGDVFFNTGDLLVCDDQGFLRFHRTGDTFRWKGENVATTEVAEVFEALDFLQEV
NVYGVTPVPGHEGRAGMAALVLRPPHALDLMQLYTHVSENLPYPARPRFLRLQESLATTETFK
QQKVRMANEGFDPSTLSDPLYVLDAQVAYLPLTTARYSALLAGNLRI

Type II transmembrane domain:

amino acids 45-65

Other transmembrane domain:

amino acids 379-398

cAMP- and cGMP-dependent protein kinase phosphorylation site
starting at amino acid 136

CUB domain protein motif

amino acids 254-261

putative AMP-binding domain signature

amino acids 332-343

N-glycosylation sites

amino acids 37-40 and 483-486

FIGURE 40

CCTGTGTTAAGCTGAGGTTTCCCCTAGATCTCGTATATCCCCAACACATACCTCCACGCACA
CACATCCCCAAGAACCTCGAGCTCACACCAACAGACACACGCGCGCATACACACTCGCTCTC
GCTTGTCCATCTCCCTCCCGGGGAGCCGGCGCGCTCCCACTTTGCCGCACACTCCGGC
GAGCCGAGCCCGCAGCGCTCCAGGATTCTGCGGCTCGGAACTCGGATTGCAGCTCTGAACCC
CCATGGTGTGTTTTTAAACACTTCTTTCTCTCTCTCTCGTTTTGATTGCAACCGTTTCCA
TCTGGGGGCTAGAGGAGCAAGGCAGCAGCCTTCCAGCCACCCCTTGTGTGGCTTGCCATCGT
CCATCTGGCTTATAAAAGTTTGCTGAGCGCAGTCCAGAGGGCTGCGCTGCTCGTCCCCTCGG
CTGGCAGAGGGGGTGACGCTGGGCAGCGGGCAGGAGCGCGCGCTCGCTCTGGCGGGCTTT
CGGCTTGAGGGGCAAGGTGAAGAGCGCACCGGCCGTGGGGTTACCGAGCTGGATTTGTATG
TTGCACCAATGCCCTTCTTGGATCGGGCTGTGATTCTTCCCCTCTTGGGGCTGCTGCTCTCCC
TCCCCCGCGGGCGGATGTGAAGGCTCGGAGCTGCGGAGAGGTCGCCAGGCGTACGGTGCC
AAGGATTACAGCTGGCGGCATCCCTACCAGGAGATCGCAGGGGAACACTTAAGAATCTG
TCCTCAGGAATATACATGCTGCACCAAGAAATGGAAGACAAGTTAAGCCAAACAAGCAAA
TCGAATTTGAAAACCTTGTGGAAGAGACAAGCCATTTTGTGCGCACCACTTTTGTGCCAGG
CATAAGAAATTTGACGAATTTTTCCGAGAGCTCCTGGAGAATGCAGAAAAGTCACTAAATGA
TATGTTGTGACGACCTATGGCATGCTGTACATGCAGAATTCAGAAGTCTTCCAGGACCTCT
TCACAGAGCTGAAAAGGTACTACACTGGGGTAATGTGAATCTGGAGGAATGCTCAATGCT
TTTTGGGCTCGGCTCCTGGAACGGATGTTTCAGCTGATAAACCTCAGTATCACTTCAGTGA
AGACTACCTGGAATGTGTGAGCAAATACACTGACCAGCTCAAGCCATTTGGAGACGTGCCCC
GGAACTGAAGATTACAGTTACCCGCGCCTTATGTGTCGACGAGCCTTTGTCCAGGGGCTG
ACTGTGGGCAGAGAAGTTGCAAAACCGAGTTTCCAAGTTCAGCCCCAACCCAGGGTGTATCCG
TGCCCTCATGAAGAATGCTGTACTGCCATACTGTGGGGGCTTCCCATGTGAGGCCCTGCA
ACAACTACTGTCTCAACGTCATGAAGGGCTGCTTGGCAAATCAGGCTGACCTCGACACAGAG
TGGAAATCTGTTTATAGATGCAATGCTCTTGGTGGCAGAGCGACTGGAGGGGCCATTCAACAT
TGAGTCGGTCATGGACCCGATAGATGTCAAGATTTCTGAAGCCATTATGAACATGCAAGAAA
ACAGCATGCAAGTGTCTGCAAAAGTCTTTCAAGGATGTGGTCAGCCCCAAACCTGCTCCAGCC
CTCAGATCTGCCCGCTCAGCTCCTGAAAATTTTAAATACAGTTTTCAGGCCCTACAATCCTGA
GGAAAGACAACAACCTGCTGCAGGCACAAGCTTGGACCGGCTGGTCACAGACATAAAAGAGA
AATTGAAGCTCTCTAAAAGGTCTGGTCAGCATTACCTTACATCTCTGCAAGGACGAGAGC
GTGACAGCGGGCAGTCCAAAGGAGGAATGCTGGAACGGGCACAGCAAAGCCAGATACCTT
GCCTGAGATCATGAATGATGGGCTCACCAACAGATCAACAATCCCGAGGTGGATGTGGACA
TCACCTCGGCCTGACACTTTTATCAGACAGCAGATATATGGCTCTCGTGTGATGACCAACAAA
CTAAAAAACCGCTTACAATGGCAATGATGTCAATTTCCAGGACACAAGTATGATGAATCCAGTGG
CTCAGGGAGTGGCAGTGGGTGCATGGATGACGTGTGTCCCACGGAGTTTGAGTTTGTCACCA
CAGAGGCCCCCGCAGTGGATCCGACCGGAGAGAGGTGGACTCTTCTGCAGCCAGCGTGGC
CACTCCCTGCTCCTGGTCTCTCACCTGCATTGTCTGGCAGCTGAGAGACTGTGCAGATGA
ATCTTTGGGTTTTTGGTCAGATGAAATCGATTTTAGCTATCTGAATGGCCAACTCACTTCTT
TTCTTACATCTTTGGACAATGGACCATGCCAAAACTTACCCTTTCTATGAGAAGAGAG
CAGTAATGCAATCTGCCTCCCTTTTTTGTTTTTCCCAAAGAGTACCGGGTGCCAGCAAGATG
CTCTCTCTTCTCTTCACTATCTGTGGGGACCTGTGTTTATTCTAGAGAGAATTCTTACTCAA
ATTTTTCGTACAGGAGATTTTCTACCTTTCATTGCTTTTATGCTGCAGAAGTAAGGAAT
CTCAGGTTGTGAGGGTTTTTTTTTTCTCATTTAAAAAT

FIGURE 41

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50914
><subunit 1 of 1, 555 aa, 1 stop
><MW: 62736, pI: 5.36, NX(S/T): 0
MPSWIGAVILPLLGLLLSLPAGADVKARSCGEVRQAYGAKGFSLADIPYQEIAGEHLRICPQ
EYTCCTTEMEDKLSQQSKLEFENLVEETSHFVRTTFVSRHKKFDEFFRELLENAEKSINDMF
VRTYGMLYMQNSEVFPQDLFTELKRYTGGNVNLEEMLNDFWARLLERMFQLINPQYHFSEDY
LECVSKYTDQLKPFQDVPRKLKIQVTRAFIAARTFVQGLTVGREVANRVSKVSPTPGCIRAL
MKMLYCPYCRGLPTVRPCNNYCLNVMGCLANQADLDTAWNLFIDAMLLVAERLEGPFNIES
VMDPIDVKISEAIMNMQENSMQVSAKVFQCGQPKPAPALRSARSAPENFNTFRFPYNPEER
PTTAAGTSLDRLVTDIKEKLKLSKKVWSALPTYICKDESVTAGTSNEEECWNGHASKARYLPE
IMNDGLTNQINNPEVDVDITRPDTFIRQQIMALRVMTNKLKNAYNQNDVNFQDTSDESSGSG
SGSGCMDDVCPTFEFEFVTTTEAPAVDPDRREVDSSAAQRGHSLLSWSLTCIVLALQRLCR
```

Signal peptide:

amino acids 1-23

FIGURE 42A

CGGACGCGTGGCGGACGCGTGGGCAAAAGAACTCGGAGTGCCAAAGCTAAATAAGTTAGCTGAGAAAAACGACG
CAGTTTGCAGCGCCTGCGCCGGGTGCGCCAACTACGCAAGACCAAGCGGGCTCCGCGCGGACCGGCCGCGGGG
TAGGGACCCGCGCTTGTGCCTTCAGGCTCCCTAGCAGCGGGGAAAGGAATGTCTGCCCGGAGTTTCTGCGGAGGT
GGAGGGAGATCAGGAAACGGCTTCTTCTCACTTCGCGCGCTTGGTGTCTCGGGGAGATTGGCAAAACGCTAGG
AAAGGACTCGGGGAAATAGCCCTGGGAAAGTGGAGAAGGTGATCAGGAGGCGGCTCCACTACGGCAGTTTATCTG
TCTGATCAGAGCCAGACGCGACGCTCCACTTTCGAGTTCTTTCAGGTGTGGGACCGCAGGACAGACGCGCGA
TCCGCGCGCCTTCGTACAGCACTCCAGGAGAGTCAGCCTCGCTCCCCAACGTCGAGGGCGCTCTGCGCACGA
AAAGTTCCTGTCCACTGTGATTCTCAATCTCTGCTTGGTTTTTTTCTCCAGAGAACTTTTGGGTGGAGATATTA
ACTTTTTCTTTTTTTTTTCTTGGTGAAGCTGCTCTAGGAGGGGGAGGAGGAGGAGAAAGTGAATGTGCT
TGGAGAAGAGCGAGCCCTCTTGTCTTCCGGAGTCCCATCCATTAAGCCATCACTTCTGGAAGATTAAAGTTGT
CGGACATGCTGACAGCTGAGAGGAGAGGAGGATTTCTTGCAGGTGGAGAGTCTTCAACGCTCTGTTGGGTGCATG
TGTGCGCCCGCAGCGGCGCGGGGCGCTGTTCTCCGCTGGAGTCTCACTGGGACCTGAGTGAATGGCTCCCA
GGGCTGTGCGGGGCATCCGCTCCGCTCTCTCCACAGGCTGTGTCTGCTCTGGAAAGATGCTAGCAATGGGG
CGCTGCGAGGATTCTGATCTCTGCTGCTCTCACTTATGTTTAACTGTCTGGGCGCAGGCTTTAGAGAGGAGG
AAGAAGGGGCTTACTAGCTCAAGCTGGAGAGAACTAGAGCCAGCACAACTTCCAGTCCCGAGCCCATCTCA
TTTTTCTCTAGCGGATGATCAGGAGTTTAGAGATGTGGTTTACCAAGGATCTGAGATTAAACACCTACTCTTG
ACAAGCTCGCTGCCAAGGAGTTAAACTGGAGAACTACTATGTCCAGCCTATTGTGACACCATCCAGGAGTCAGT
TTATTACTGGAAGTATCAGATACACCCGGACTTCAACATTTATCATAAGACCTACCCAAACCACTGTTTAC
CTCTGGACAATGACCACCCTACCTCAGAAACTGAAGGAGTTGGATATTCAACGCATATGTCGGGAAATGGCACT
TGGGTTTTAACAGAAAAGATGCATGCCCAACAGAGAGGATTTGATACCTTTTTTGGTTCCCTTTTGGGAAGTG
GGGATTACTATACACTACAATGTGACAGTCTTGGGATGTGGCTATGACTTGTATGAAACGACAATGCTG
CCTGGGACTATGACAATGGCATATATCCACACAGATGTACACTCAGAGAGTACAGCAAACTTTAGCTTCCCAT
ACCCCAAAAGCCTATATTTTTATATATCTGCTTCAAGCTGTTCATTCAACCTGCAAGCTCCTGGCAGGTATT
TCGAACACTACCGATCCATTTATCAACATAAACAGGAGAAGATATGCTCCCATGCTTTCTGCTTAGATGAAGCAA
TCAACACAGTGACATTGGCTCTAAAGACTTATGGTTTCTATAACAACAGCATTATCATTTACTCTTCAGATAATG
GTGGCCAGCCTACCGCAGGAGGAGTAACTGGCCTCTCAGAGGTAGCAAGGAACATATTGGGAAGGAGGGATCC
GGGCTGTAGGCTTTGTGCATAGCCCACTTCTGAAAAACAAGGGAACAGTGTGTAAGGAACCTGTGACATCACTG
ACTGTTACCCACTCTCATTTCACTGGCTGAAGGACAGATTGATGAGGACATTCACATAGATGATGATATCT
GGGAGACCATAAGTGAGGGTCTTTCGCTCACCCCGAGTAGATATTTGCATAACATTGACCCCTATACACCAAGGC
AAAAAATGGCTCTGGGCAGCAGGCTATGGGATCTGGAACTGCAATCCAGTCAGGCCATCAGAGTGCAGCACTG
GAAATTGCTTACAGGAAATCCTGGCTACAGCACTGGTCCCCCTCAGTCTTTTCAGCAACCTGGGACCGAACC
GTGGCAATGAACCGGATCACTGTGTCACTGGCAAAAGTATATGGCTTTCAACATTCACGCGCCACCATATGA
GAGGTGGACCTATCTAACAGGTATCCAGGAATCTGAGAGAGCTCCTACGGAGGCTCTCAAGTCTCAACAAAC
TGCAGTGCCGTCAGGTATCCCCCAAAGACCCAGAAAGTAACCTAGGCTCAATGGAGGGGTCTGGGGACCATG
GTATAAAGAGGAAACCAAGAAAAGAGCCAGCAAAAATCAGGCTGAGAAAAGCAAAGAAAAGCAAAAAA
GAAGAAGAAACAGCAGAAAGCAGTCTCAGTAAACAGCAAAATTTGGCTCGATAATATCGCTGGCCTAAGCGTCA
GGCTTGTTTTCTGCTGTGCCCCTCAGAGACTTCTGCCACCTGGCCGCCACACTGAAACCTGTCTGCTCAGTG
CCAAGTGCTACTCTTGCAAGCCACACTTAGAGAGAGTGGAGATGTTTATTCTCTCGCTCCTTTAGAAAAAGTG
GTGAGTCTGAGTTTCACTGTCTGTGCTTCACTCACTGACCAAACTGCTTTGAATTATAGGAGGAGAACATA
ACCTACCATCCGCAAGCATGCTAATTTGATGGAAGTTACAGGATAGCATGATTAACACTACCTTTGATAAATTAC

CGGACGCGTGGCGGACGCGTGGGCAAAAGAACTCGGAGTGCCAAAGCTAAATAAGTTAGCTGAGAAAAACGACG

[illegible][illegible]

FIGURE 43

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48296
><subunit 1 of 1, 515 aa, 1 stop
><MW: 56885, pI: 6.49, NX(S/T): 5
MAPRGCAGHPPPPSPQACVCPGKMLAMGALAGFWILCLLTYGYLSWGQALEEEEEEGALLAQA
GEKLEPSTTSTSQPHLIFILADDQGFRDVGYPHGSEIKTPTLDKLAAGVKLENYYVQFICTP
SRSQFITGKYQIHTGLQHSIIIRPTQFNCLPLDNATLPQKLKEVGYSTHMGVKGWHLGFNRKEC
MPTRRGFDTFFGSLLGSGDYTHYKCDSPGMCGYDLYENDNAAWDYDNGIYSTQMYTQRVQQ
ILASHNPTKPIFLYTAYQAVHSPQAPGRYFEHYRSIININRRRYAAMLSCLEAINNVTLA
LKTYGFYNNSSIIYYSSDNGGQPTAGGSNWPLRSGKTYWEGGIRAVGFVHSPLLNKGTVCCK
ELVHITDWPYTLISLAEGQIDEDIQLDGYDIWETISEGLRSRPRVDILHNIDPYTPRQKMAPG
QQAMSGTLQSSQPSECTGNCLQEILATATGSPLSLSATWDRGTGGTMNGSPCQLAKVYGFSS
TSQPTHMRGWYTLTGIQES

Important Features:

Signal Peptide:

amino acids 1-37

Sulfatases signature 1.

amino acids 120-132

Sulfatases signature 2.

amino acids 168-177

Tyrosine kinase phosphorylation site.

amino acids 163-169

N-glycosylation sites.

amino acids 157-160, 306-309 and 318-321

FIGURE 44

CGGACGCGTGGGTGCGAGTGGAGCGGAGGACCCGAGCGGCTGAGGAGAGAGGAGGCGGCGGC
TTAGCTGCTACGGGGTCCGGCCGGCGCCCTCCCGAGGGGGGCTCAGGAGGAGGAAGGAGGAC
CCGTGCGAGAATGCCTCTGCCCTGGAGCCTTGCGCTCCCGTGTGCTCTCTGGTGGCAG
GTGGTTTCGGGAAACGGCGCCAGTGAAGGCATCACGGGTGTGTAGCATCGGCACGTGAGCCT
GGGGTCTGTCACTATGGAACATAAACTGGCCTGCTGCTACGGCTGGAGAAGAAACAGCAAGGG
AGTCTGTGAAGCTACATGCGAACCTGGATGTAAGTTTGGTGAAGTGGTGGGACCAACAAAT
GCAGATGCTTTCCAGGATACACCGGGAACCTGCAGTCAAGATGTGAATGAGTGTGGAATG
AAACCCCGCCATGCCAACACAGATGTGTGAATACACCGGAAGCTACAAGTGCTTTTGCTC
CAGTGGCCACATGCTCATGCCAGATGCTACGTGTGTGAACCTCTAGGACATGTGCCATGATAA
ACTGTCAGTACAGCTGTGAAGACACAGAAGAAGGGCCACAGTGCCTGTGTCCATCCTCAGGA
CTCCGCTGGCCCCAATGGAAGAGACTGTCTAGATATTGATGAATGTGCCTCTGGTAAAGT
CATCTGTCCCTACAATCGAAGATGTGTGAACACATTTGGAAGCTACTACTGCAAAATGTCA
TTGGTTTCGAACGTGCAATATATCAGTGGACGATATGACTGTATAGATATAAATGAATGTACT
ATGGATAGCCATACGTGCAGCCACCATGCCAATTGCTTCAATACCCAAGGGTCTTCAAGTG
TAAATGCAAGCAGGGATATAAAGGCAATGGACTTCGGTGTCTGCTATCCCTGAAAATTCTG
TGAAGGAAGTCTCAGAGCACCTGGTACCATCAAAGACAGAATCAAGAAGTTGCTTGTCTCAC
AAAAACAGCATGAAAAAGAAGGCCAAAAATTAAAAATGTTACCCAGAACCCACCAGGACTCC
TACCCTAAGGTGAACCTTGCAGCCCTTCAACTATGAAGAGATAGTTTCCAGAGGCGGGAAC
CTCATGGAGGTAAAAAAGGGAATGAAGAGAAATGAAGAGGGGCTTGAGGATGAGAAAAGAG
AAGAGAAAGCCCTGAAGAAATGACATAGAGGAGCGAAGCCTGCGAGGAGATGTGTTTTCCT
AAGGTGAATGAAGCAGGTGAATTCGGCCTGATTCTGGTCCAAAGGAAAGCGCTAACTTCCAA
ACTGGAACATAAAGATTAAATATCTCGGTTGACTGCAGCTTCAATCATGGGATCTGTGACT
GGAAACAGGATAGAGAAGATGATTTTGACTGGAATCCTGCTGATCGAGATAATGCTATTGGC
TTCTATATGGCAGTTCCGGCCTTGGCAGGTCAACAAGAAAGACATTGGCCGATTGAAACTTCT
CCTACCTGACCTGCAACCCCAAGCAACTTCTGTTTGTCTTTTGATTACCGGCTGGCCGGAG
ACAAAGTCGGGAAACTTCGAGTGTTGTGAAAAACAGTAAACAATGCCCTGGCATGGGAGAAG
ACCACGAGTGAGGATGAAAAGTGAAGACAGGGAATAATCAGTTGTATCAAGGAACATGATGC
TACCAAAAGCATCATTTTTGAAGCAGAACGTGGCAAGGGCAAAACCGGCGAAATCGCAGTGG
ATGGCGTCTTGCTGTTTTCAGGCTTATGTCCAGATAGCCTTTTATCTGTGGATGACTGAATG
TTACTATCTTTATATTGACTTTGTATGTGAGTCCCTGGTTTTTTTGATATTGCATCATAG
GACCTCTGGCATTTTAGAATTACTAGCTGAAAAATTGTAATGTACCAACAGAAATATTATTG
TAAGATGCCCTTCTGTATAAGATATGCCAATATTGCTTTAAATATCATATCACTGTATCT
TCTCAGTCATTTCTGAATCTTCCNCATTATATTATAAAATNTGGAAANGTCAGTTTATCTC
CCCTCCTCNGTATATCTGATTGTATANGTANGTTGATGNGCTTCTCTACAACTTTCTA
GAAATAGAAAAAAGCACAGAGAAATGTTTAACTGTTTGACTCTTATGATACTTCTTGA
AACTATGACATCAAAAGATAGACTTTTGCCCTAAGTGGCTTAGCTGGGTCTTTCATAGCCAAAC
TTGTATATTAAATCTTTGTAATAATAA

FIGURE 45

MPLPWSLALPLLLSWAGGFGNAASARHHGLLASARQPGVCHYGTKLACCYGWRRNSKGVCE
ATCEPGCKFGECVGPNNKRCFPGYTGKTCSDVNECGMKPRPCQHRVCVNTHGSIKCFCLSGH
MLMPDATCVNSRTCAMINCQYSCEDTEEGPQCLCPSSGLRLAPNGRDCLDIDECASGKVICP
YNRRVCNTFGSIYCKCHIGFELQYISGRYDCIDINECTMDSHTCSHHANCFNTQGSFKCKCK
QGYKGNGLRCSAIPENSVKEVLRAPGTIKDRIKKLLAHKNSMKKKAKIKNVTPEPTRTPTPK
VNLQPFNYEEIVSRGGNSHGGKKGNEEK

Signal peptide:

amino acids 1-21

EGF-like domain cysteine pattern signature.

amino acids 80-91

Calcium-binding EGF-like domains

amino acids 103-124, 230-251 and 185-206

FIGURE 46

GGGAGCTGCTGCTGTGGCTGCTGGTGCTGTGCGCGCTGCTCCTGCTCTTGGTGCAGCTGCTG
CGCTTCTGAGGGCTGACGGCGACCTGACGCTACTATGGGCCGAGTGGCAGGGACGACGCC
AGAATGGGAGCTGACTGATATGGTGGTGTGGGTGACTGGAGCCTCGAGTGGAATTGGTGAGG
AGCTGGCTTACCAGTTGTCTAACTAGGAGTTTCTCTTGTGCTGTCAGCCAGAAGAGTGCAT
GAGCTGGAAGGGTGAAGAAGATGCC TAGAGAATGGCAATTTAAAGAAAAAGATATACT
TGTTTGGCCCTTGACCTGACCGACACTGGTTC CATGAAGCGGCTACCAAAGCTGTTCTCC
AGGAGTTGGTAGAATCGACATTCTGGTCAACAATGGTGGAATGTCCAGCGTTCTCTGTGC
ATGGATACCAGCTTGATGTCTACAGAAAGCTAATAGAGCTTAACTACTTAGGGACGGTGTG
CTTGACAAAATGTGTTCTGCCTCACATGATCGAGAGGAAGCAAGGAAAGATTGTTACTGTGA
ATAGCATCTGGGTATCATATCTGTACCTCTTCCATTGGATACTGTGCTAGCAAGCATGCT
CTCCGGGGTTTTTTTAAATGGCCTTCGAACAGAACTTGCCACATACCCAGGTATAATAGTTTC
TAACATTTGCCCAGGACCTGTGCAATCAAATATTGTGGAGAATTCCTTAGCTGGAGAAGTCA
CAAAGACTATAGGCAATAATGGAGACCAGTCCCAAGATGACAACCAAGTCGTTGTGTGCGG
CTGATGTTAATCAGCATGGCCAATGATTTGAAAGAAGTTGGATCTCAGAACCAACCTTCTTT
GTTAGTAACATATTTGTGGCAATACATGCCAACCTGGGCTGGTGGATAACCAACAAGATGG
GGAAGAAAAGGATTGAGAACTTTAAGAGTGGTGTGGATGCAGACTCTTCTTATTTTAAATC
TTTAAGACAAAACATTGACTGAAAAGAGCACCTGTACTTTTCAAGCCACTGGAGGGAGAAATG
GAAAACATGAAAACAGCAATCTTCTTATGCTTCTGAATAATCAAGACTAATTTGTGATTTT
ACTTTTAAATAGATATGACTTTGCTTCCAACATGGAATGAAATAAAAAATAAATAATAAAAG
ATTGCCATGAATCTTGCAAAA

FIGURE 47

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA36343
><subunit 1 of 1, 289 aa, 1 stop
><MW: 32268, pI: 9.21, NX(S/T): 0
MVVVVTGASSGIGEELAYQLSKLGVSLVLSARRVHELERVKRRCLENGNLKEKDILVLPLDL
TDTGSHEAATKAVLQEFGRIDILVNNGGMSQRSCLCMDTSLDVYRKLIELNYLGTVSLTKCVL
PHMIERKQGKIVTVNSILGIISVPLSIGYCASKHALRGFFNGLRTELATYPGIIVSNICPGP
VQSNIVENSLAGEVTKTIGNNGDQSHKMTTSRCVRLMLISMANDLKEVWISEQPFLLVTYLW
QYMPTWAWWITNKMGGKKRIENFKSGVDADSSYFKIFKTKHD

Important Features:

Signal Peptide:

amino acids 1-31

Transmembrane domain:

amino acids 136-157

Tyrosine kinase phosphorylation site.

106-113 and 107-114

Homologous region to Short-chain alcohol dehydrogenase

amino acids 80-90, 131-168, 1-13 and 176-185

FIGURE 48

GCGACGTGGGCACCGCCATCAGCTGTTGCGCGCTCTTCTCCTCCAGGTGGGGCAGGGGTTTC
GGGCTGGTGGAGCATGTGCTGGGACAGGACAGCATCCTCAATCAATCCAACAGCATATTCCGC
TTGCATCTTCTACACACTACAGCTATTGTTAGGTTGCCTGCGGACACGCTGGGCCCTCTGTCC
TGATGCTGCTGAGCTCCCTGGTGTCTCTCGCTGGTTCTGTCTACCTGGCCTGGATCCAGTCCG
TTCGTGCTCTATGATTCTTGCATTGTTTGTATCACCACCTATGCTATCAACGTGAGCCTGAT
GTGGCTCAGTTTCCGGAAGGTCCAAGAACCCAGGGCAAGGCTAAGAGGCACTGAGCCCTCA
ACCCAAGCCAGGCTGACCTCATCTGCTTTGCTTTGGTCTTCAAGCCGCTCAGCGTGCCTGTG
GACAGCGTGGCCCCGGCCCCCAGCCCTCAGGAGGGCAACACAGTCCCTGGCGAGTGGCCCC
TGGCAGGCCAGTGTGAGGAGGCAAGGAGCCCCACATCTGCAGCGGCTCCCTGGTGGCAGACAC
CTGGGTCTCTCACTGCTGCCCCTGCTTTGAAAAGGCAGCAGCAACAGAAGTGAATTCCTGGT
CAGTGGTCTCTGGGTTCTCTGCAGCGTGAGGGACTCAGCCCTGGGGCCGAAGAGGTGGGGGTG
GCTGCCCCGTCAGTTGCCAGGGCCATAAACCTACAGCCAGGGCTCAGACCTGGCCCTGCT
GCAGCTCGCCCCACCCACGACCCACACACCCCTCTGCTGCCCCAGCCCCGCCCATGCTTCC
CCTTTGGAGCCTCTGCTGGGCCACTGGCTGGGATCAGGACACAGTGATGCTCCTGGGACC
CTACGCAATCTGGCCCTGCGTCTCATCAGTCGCCCCACATGTAAGTGTATCTACAACAGCT
GCACCAGCGACACCTGTCCAACCCGGCCCGGCTGGGATGCTATGTGGGGGCCCCAGCCTG
GGGTGCAGGGCCCCCTGTGAGGGAGATTCCGGGGGCCCTGTGCTGTGCCCTCAGACCTGACGGA
CACTGGGTTCAAGCTGGCATCATCAGCTTTGCATCAAGCTGTGCCAGGAGGACGCTCCTGT
GCTGTGACCAACACAGCTGCTCAGGTTCTCTGGCTGCAGGCTCGAGTTCAAGGGGCGAGCTT
TCCTGGCCCCAGAGCCAGAGACCCCGAGATGAGTGATGAGGACAGCTGTGTAGCCTGTGGA
TCCTTGAGGACAGCAGGTTCCCCAGGCAGGAGCACCTCCCATAGGCCCTGGGAGGCCAGGCT
GATGCACCCAGGGACAGCTGGCCTGTGGCGGAGCCCTGGTGTGAGGAGGCCGTGCTAACTG
CTGCCCACTGCTTCATTGGGCGCCAGGCCCCAGAGGAATGGAGCGTAGGGCTGGGGACCGA
CCGGAGGAGTGGGGCTGGAAGCAGCTCATCTGTCATGGAGCCTACACCCACCTGAGGGGGG
CTACGACATGGCCCTCTGCTGCTGGCCCCAGCCTGTGACACTGGGAGCCAGCCTGCGGCCCC
TCTGCTGCCCCATCTCTGACCAACACCTGCTGATGGGAGCGTGGCTGGGTTCTGGGACCG
GCCGCCAGGAGCAGGCATCAGCTCCCTCCAGACAGTGCCCGTGACCCTCCTGGGGCCTAG
GGCCTGCAGCCCGCTGCATGACAGCTCCTGGGGGTGATGGCAGCCCTATTCTGCCGGGGATGG
TGTGTACCAAGTGTGTGGGTGAGCTGCCAGCTGTGAGGGCCTGTCTGGGGCACCACTGGTG
CATGAGGTGAGGGGCACATGGTTCTCTGCGCCGGCTGCACAGCTCTGGAGATGCTTGCCAAG
CCCCGCCAGGCCCGCGGTCTTACCGCGCTCCTGCTATGAGGACTGGGTGAGCAGTTTGG
ACTGGCAGGTCTACTTTCGCCGAGGAACAGAGCCCGAGGCTGAGCCTGGAAGCTGCCTGGCC
AACATAAGCCAACCAACAGCTGCTGACAGGGGACCTGGCCATTCTCAGGACAAGAGAATGC
AGGCAGGCCAATGGCAATTACTGCCCCCTGTCTCCCCACCTGTGATGTGTGATTCCAGGCAC
CAGGCAGGCCCAGAAGCCAGCAGCTGTGGGAAGGAACCTGCTGGGGCCACAGGTGCCCA
CTCCCCACCTGCAGGACAGGGGTGTCTGTGGACACTCCACACCCAACTCTGCTACCAAGC
AGGCTGTCTCAGCTTTCTCTCTCTTACTCTTTTCAGATACAATCAGCCAGCCACGTTGTTT
TGAAAATTTCTTTTTTGGGGGCGCAGGTTTTCTTTTTTAACTTAAATAAATTTGTTAC
AAAAATAAA

FIGURE 49

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA40571

MLLSSSLVSLAGSVYLAWILFFVLYDFCIVCITTYAINVSLMWLSFRKVQEPQKAKRHGNTV
PGEWPWQASVRRQGAHICSGSLVADTWVLTAAHCFEKAAATELNSWSVVLGSLQREGLSPGA
EEVGVAAQLPRAYNHYSQGSDDLALLQLAHPTTHTPLCLPQPAHRFPFGASCWATGWDQDTS
DAPGTLRNLRLRLISRPTCNCIYNQLHQRHLSNPARPGMLCGGPQPGVQGPCQGDSGGPVLC
LEPDGHWVQAGIISFASSCAQEDAPVLLTNTAAHSSWLQARVQGAAFLAQSPETPEMSDEDS
CVACGSLRTAGPQAGAPSPWPWEARLMHQQLACGGALVSEEAVLTAHCFIGRQAPEEWSV
GLGTRPEEWGLKQLILHGAYTHPEGGYDMALLLLAQPVTLGASLRPLCLPYPDHHLDPDGERG
WVLGRARPGAGISSLQTVPVVTLGPRACSRLLHAAPGGDGSPILPGMVCTSAVGELPSCEGLS
GAPLVHEVRGTWFLAGLHSFGDACQGPAPPAVFTALPAYEDWVSSLDWQVYFAEEPEPEAEF
GSCLANISQPTSC

Important features:

Signal peptide:

amino acids 1-15

Homologous region to Serine proteases, trypsin family

amino acids 79-95, 343-359 and 237-247

N-glycosylation sites.

amino acids 37-40 and 564-567

Kringle domains

amino acids 79-96, 343-360 and 235-247

FIGURE 50

CGGGCCGCCCGGCCCCCATTCGGGCCGGGCCTCGCTGCGGCGGCGACTGAGCCAGGCTGG
GCCGCGTCCCTGAGTCCCAGAGTCGGCGCGGCGCGGAGGGGAGCCTTCCACCACGGGGAG
CCCAGCTGTCCAGCCCTCACAGGAAGATGCTGCGTCGGCGGGGAGCCCTGGCATGGGTGT
GCATGTGGGTGCAGCCCTGGGAGCACTGTGGTTCTGCCTCACAGGAGCCCTGGAGGTCCAGG
TCCCTGAAGACCCAGTGGTGGCACTGGTGGGCACCGATGCCACCTGTGCTGCTCCTTCTCC
CCTGAGCCTGGCTTCAGCCTGGCACAGCTCAACCTCATCTGGCAGCTGACAGATACCAAACA
GCTGTGTCACAGCTTTGCTGAGGGCCAGGACCAGGGCAGCGCTATGCCAACGACGCGCCC
TCTTCCCGGACCTGCTGGCACAGGGCAACGCATCCCTGAGGCTGCAGCGCTGCGTGTGGCG
GACGAGGGCAGCTTCACCTGCTTCGTGAGCATCCGGGATTTGCGGAGCGCTGCCGTGAGCCT
GCAGTGGCCGCTCCCTACTCGAAGCCCAGCATGACCTGGAGCCCCAACAGGACCTGCGGC
CAGGGGACACGGTGACCATCACGTGCTCCAGCTACCAGGGCTACCTGAGGCTGAGGTGTTT
TGGCAGGATGGGCGAGGTGTGCCCTGACTGGCAACGTGACCACGTGCGAGATGGCCAACGA
GCAGGGCTTGTGTTGATGTGCACAGCGTCTGCGGGTGGTGTGGGTGCGAATGGCACCTACA
GCTGCCTGGTGCACAACCCGTGCTGCAGCAGGATGCGCACRGTCTGTCAACATCACAGGG
CAGCCTATGACATTTCCCCCAGAGGCCCTGTGGGTGACCGTGGGGCTGTCTGTCTGTCTCAT
TGCACTGCTGGTGCCCTGGCTTTCTGTGTCTGGAGAAAGATCAAACAGAGCTGTGAGGAGG
AGAAATGCAGGAGTGTAGGACCAGGATGGGGAGGGAGAAGGCTCCAAGACAGCCCTGCAGCCT
CTGAAACACTCTGACAGCAAGAAGATGATGGACAAGAAATAGCCTGACCATGAGGACCAGG
GAGTGCTACCCCTCCCTACAGCTCCTACCCCTGGCTGCAATGGGGCTGCACTGTGAGCCC
TGCCCCAACAGATGCATCCTGCTCTGACAGGTGGGCTCCTTCTCCAAGGATGCGATACAC
AGACCACTGTGCAGCCTTATTTCTCCAATGGACATGATCCCAAGTCATCCTGCTGCCTTTT
TCTTATAGACACAATGAACAGACCACCCACAACCTTAGTTCTCTAAGTCATCCTGCTGCT
GCCTTATTTACAGTACATACATTTCTTAGGGACACAGTACACTGACCACATACCAACCTC
TCTTCCAGTGTGCTGGTGACCATCTGGCTGCCTTTTCTCCAAAAGATGCAATATTGAGA
CTGACTGACCCCTGCGCTTATTTACCAAAGACACGATGCATAGTCACCCGCGCTTGTTC
TCCAATGGCCGTGATACACTAGTGATCATGTTTCAGCCCTGCTTCCACCTGCATAGAATCTTT
TCTTCTCAGACAGGGACAGTGC GGCTCAACATCTCCTGGAGTCTAGAAGCTGTTTCTTTT
CCCTCCTTCTCCCTGCCCAAGTGAAGACAGGGCAGGGCCAGGAATGCTTTGGGGACACCG
AGGGGACTGCCCCCACCACCATGGTGCTATTCTGGGGCTGGGGCAGTCTTTTCTGGC
TTGCTCTGGCCAGCTCCTGGCCTCTGGTAGAGTGAGACTTCAGACGTTCTGATGCCTTCCG
GATGTCATCTCTCCCTGCCCCAGGAATGGAAGATGTGAGGACTTCTAATTTAAATGTGGGAC
TCGGAGGGATTTGTAAACTGGGGGTATATTTTGGGGAATAAATGTCTTTGTAAAAAAA
AAAAAAAAAAAAA

FIGURE 51

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA41386
><subunit 1 of 1, 316 aa, 1 stop, 1 unknown
><MW: -1, pI: 4.62, NX(S/T): 4
MLRRRGSPGMGVHVGAAALGALWFCLTGALEVQVPEDPVVALVGTDATLCCSFSPPEPGFSLAQ
LNLIWQLTDTKQLVHSFAEGQDQGSAYANRTALFPDLLAQGNASRLRQVRVVADEGSFTCFV
SIRDFGSAAVSLQVAAPYSKPSMTLEPNKDLRPGDTVTITCSSYQGYPEAEVFWQDQGQGVPL
TGNVTTSQMANEQGLFDVHSLRVVLGANGTYSCLVRNPVLQQDAHXSVTITGQPMTFPPEA
LWVTVGLSVCLIALLVALAFVCWRKIKQSCEEENAGAEDQDGELEGSKTALQPLKHSDSKED
DGQEIA

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 251-270

N-glycosylation site.

amino acids 91-94, 104-107, 189-192 and 215-218

Homologous region to Immunoglobulins and MHC

amino acids 217-234

FIGURE 52

TTCGTGACCCCTTGAGAAAAAGAGTTGGTGGTAAATGTGCCACGCTCTCTAAGAAGGGGGAGTC
CTGAACCTTGCTGAAGCCCTTGTCGCTAAGCCTTGAACTACGTTCTTAAATCTATGAAGTCG
AGGGACCTTTCGCTGCTTTTGTAGGGACTTCTTTCCTTGCTTTCAGCAACATGAGGCTTTTCT
TGTGGAAACGCGGTCTTGACTCTGTTTCGCTCACTTCTTTGATTGGGGCTTTGATCCCTGAACCA
GAAGTGAAAATTGAAGTTCTCCAGAACCCATTTCATCTGCCATCGCAAGACCAAAGGAGGGGA
TTTGATGTTGGTCCACTATGAAGGCTACTTAGAAAAGGACGGCTCCTTATTTCACCTCCACTC
ACAAACATAACAATGGTCAGCCCATTTGGTTTACCCTGGGCATCTCGGAGGCTCTCAAAGGT
TGGGACCAGGGCTTGAAGAGGAATGTTGTGTAGGAGAGAGAAGAAAGCTCATCATTCCTCCTGCT
TCTGGGCTATGGAAAAGAAGGAAAAGGTAAAAATTCCCCCAGAAAGTACACTGATATTTTAATA
TTGATCTCCTGGAGATTGCAAAATGGAACCAAGATCCCATGAATCATTTCCAAGAAATGGATCTT
AATGATGACTGGAACCTCTCTAAAGATGAGGTTAAAGCATATTTAAAGAAGGAGTTTGAAAA
ACATGGTGCGGTGGTGAATGAAAGTCATCATGATGCTTTGGTGGAGGATATTTTGTATAAAG
AAGATGAAGACAAAGATGGGTTTATATCTGCCAGAGAATTTACATATAAACACGATGAGTTA
TAGAGATACACTTACCCCTTTTAAATAGCACTCATCTTTCAAGAGAGGGCAGTCATCTTAA
AGAACATTTTATTTTTATACAATGTTCTTTCTTGCTTTGTTTTTATTTTTATATATTTTTT
CTGACTCCTATTTAAAGAACCCCTTAGGTTTCTAAGTACCCATTTCTTTCTGATAAGTTATT
GGGAAGAAAAAGCTAATTGGTCTTTGAATAGAAGACTTCTGGACAATTTTTCACTTTTCACAG
ATATGAAGCTTTGTTTTACTTTCTCACTTATAAATTTAAAAATGTTGCACTGGGAATATACC
ACGACATGAGACCAGGTTATAGCACAAAATAGCACCCCTATATTTCTGCTTCCCTCTATTTTT
TCCAAGTTAGAGGTCAACATTTGAAAAGCCTTTTGCAATAGCCCCAAGGCTTGCTATTTTTCAT
GTTATAATGAAATAGTTTATGTGTAAGTGGCTCTGAGTCTCTGCTTGAGGACCAGAGGAAAA
TGGTTGTTGGACCTGACTTGTGTTAATGGCTACTGCTTTACTAAGGAGATGTGCAATGCTGAAG
TTAGAAAACAGGTTAATAGCCAGGCATGGTGGCTCATGCTCTGAATCCAGCACTTTGCGGAG
GCTGAGGCGGGCGGATCACCTGAGGTTGGGAGTTCGAGACCAGCCTGACCAACACGGGAGAA
CCCTATCTCTACTAAAAATACAAAGTAGCCCGCGTGGTGATGCGTGCCCTGTAATCCAGCT
ACCCAGGAAGGCTGAGGCGGCAGAATCACTTGAACCCGAGGCGGAGGTTGCGGTAAGCCGAG
ATCACCTNCAGCTGGACACTCTGTCTCGAAAAAAGAAAAGAACCGGTTAATACCATATNA
ATATGTATGCATTGAGACATGCTACCTAGGACTTAAGCTGATGAAGCTTGGCTCTCTAGTGAT
TGGTGGCCCTATTATGATAAATAGGACAAATCATTATATGTGTGAGTTTCTTTGTAATAAAATG
TATCAATATGTTATAGATGAGGTAGAAAGTTATATTTATATTTCAATATTTACTTCTTAAGGC
TAGCGGAATATCCTTCTCGGTTCTTTAATGGGTAGTCTATAGTATATTACTTCAATAAACA
TTGTATCATAAAGATAAAGTAGTAAACAGGCTACATTTTCCCATTCTGTCTCATCAAAAA
TGAACTTAGCTGGGTGTGGTGGCTCATGCCTGTAATCCAGCACTTTGGGGGCCAAGGAGGG
TGGATCACTTGAGATCAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAAACCTTGTCTCTA
CTAAAAATACAAAAATTAGCCAGGCGTGGTGGTGACACCTGTAGTCCCAGCTACTCGGGAG
GCTGAGACAGGAGATTGCTTTGAACCCGGGAGGCGGAGGTTGCAGTGAGCCAAGATTGTGCC
ACTGCACTCCAGCCTGGGTGACAGAGCAAGACTCCATCTCAAAAAAAGAAAGCAGAG
CCTACAGCAGCTACTATTGAATAAATACCTATCCTGGATTTT

FIGURE 53

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44194
><subunit 1 of 1, 211 aa, 1 stop
><MW: 24172, pI: 5.99, NX(S/T): 1
MRLFLWNAVLTLFVTSLIGALIPePEVKIEVLQKPFICHRKTKGGDLMLVHYEGYLEKDGS
L
FHSTHKHNNQPIWFTLIGILEALKGWDQGLKGMCVGEEKRKLIIPPALGYGKEGKGKIPPE
ST
LIFNIDLLEIRNGPRSHESFQEMDLNDDWKLKDEVKAYLKKEFEKKGAVVNESHHDALVED
IFDKEDDKDGFISAREFTYKHDEL
```

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 176-179

Casein kinase II phosphorylation site.

amino acids 143-146, 156-159, 178-181 and 200-203

Endoplasmic reticulum targeting sequence.

amino acids 208-211

FKBP-type peptidyl-prolyl cis-trans isomerase

amino acids 78-114 and 118-131

EF-hand calcium-binding domain.

amino acids 191-203, 184-203 and 140-159

S-100/ICaBP type calcium binding domain

amino acids 183-203

FIGURE 54

AATAAAGCTTCCTTAATGTTGTATATGTCCTTTGAAGTACATCCGTGCATTTTTTTTTAGCAT
CCAACCATTCCTCCCTTGTAGTTCTCGCCCCCTCAAATCACCCCTCTCCCGTAGCCCCACCCGA
CTAACATCTCAGTCTCTGAAAATGCAAGAGATGCTGGCTACCTCGCCCTGCCTTCAGCCT
CACGGGGCTCAGTCTCTTTTTCTCTTTGGTGCCACCAGGACGGAGCATGGAGGTACAGTAC
CTGCCACCTCAACGTCTCAATGGCTCTGACGCCGCCCTGCCCTGCACCTTCAACTCCTCTG
TACACAGTGAACCACAACAGTTCTCCCTGAACTGGACTTACCAGGAGTGCAACAACCTGCTC
TGAGGAGATGTTCTCCAGTTCGGCATGAAGATCATTAACCTGAAGCTGGAGCGGTTTCAAG
ACCGCGTGGAGTTCTCAGGGAACCCAGCAAGTACGATGTGTCGGTGATGCTGAGAAAACGTG
CAGCCGGAGGATGAGGGGATTTACAACCTGCTACATCATGAACCCCTTGACGCCACCGCTGG
CCATGGCAAGATCCATCTGCAGGTCTCATGGAAGAGCCCCCTGAGCGGGACTCCACGGTGG
CCGTGATTGTGGGTGCCCTCCGTCGGGGGCTTCTGGCTGTGGTTCATCTTGGTGCTGATGGTG
GTCAAGTGTGTGAGGAGAAAAAAGAGCAGAAGCTGAGCACAGATGACCTGAAGACCGAGGA
GGAGGGCAAGACGGACGGTGAAGGCAACCCGGATGATGGCGCCAAGTAGTGGGTGGCCGGCC
CTGCAGCCTCCCGTGTCCGTCCTCCCTCTCCGCTGTACAGTGACCTTGCCTGCTCG
CTCTTGGTGTGCTTCCCGTGACCTAGGACCCAGGGGCCACCTGGGGCTCCTGAAACCCCG
ACTTCGTATCTCCACCTGCACCAAGAGTGACCCACTCTCTTCCATCCGAGAAACCTGCCA
TGCTCTGGGACGTGTGGGCCCTGGGGAGAGGAGAGAAAGGGCTCCACCTGCCAGTCCCTGG
GGGGAGGCAGGAGGCACATGTGAGGGTCCCCAGAGAGAAGGGAGTGGGTGGGCAGGGGTAGA
GGAGGGGCCGCTGTACCTGCCCCAGTGTCTGGCTGGCAGTGGCTTCAGAGAGGACCTGGTGG
GGAGGGAGGGCTTTCGTGTGTGACAGCGCTCCCTCAGGAGGGCTTGGCTTGGCACGGCTG
TGCTCCTCCCTGCTCCAGCCAGAGCAGCCATCAGGCTGGAGGTGACGATGAGTTCTGA
AACTTGGAGGGGCATGTTAAAGGGATGACTGTGCATTCAGGGCACTGACGGAAGCCAGGC
CTGCAGGCAAAGCTGGACATGTGCCCTGGCCAGGAGGCCATGTTGGGCCCTCGTTTCCATT
GCTAGTGGCTCCTTGGGGCTCCTGTTGGCTCCTAATCCCTTAGGACTGTGGATGAGGCCAG
ACTGGAAGAGCAGCTCCAGGTAGGGGGCCATGTTTCCAGCGGGGACCCACCAACAGAGGCC
AGTTTCAAAGTCAGCTGAGGGGCTGAGGGGTGGGGCTTCATGGTGAATGCAGTTGCTGCAG
GCTCTGCTTCTCCATGGGGTAACACCCCTCGCTGGGCAGGGGCAGCCAGGCTGGGAAAT
GAGGAGGCCATGCACAGGGTGGGGCAGCTTTCCTTGGGGCTTCAGTGAGAACTCTCCAGTT
GCCCTTGGTGGGGTTTCCACCTGGCTTTTGGCTACAGAGAGGGAAGGGAAAGCTGAGGCCG
GCATAAGGGGAGCCCTTGGAACTGAGCTGCCAATGCCAGCCCTGTCCATCTCGCGGCCAGC
CTACTCGCTCCTCTCCCAACAACCTCCCTTCGTGGGGACAAAAGTGACAATTGTAGGCCAGGC
ACAGTGGCTCACGCCCTGTAATCCCAGCACTTTGGGAGGCCAAGCGGGGTGGATTACCTCCAT
CTGTTTAGTAGAAATGGGCAAAACCCATCTCTACTAAAAATACAAGAATTAGCTGGGGCTG
GTGGCGTGTGCTGTAATCCAGCTATTGGGAGGCTGAGGCAGGAGAATCGCTTGAGCCCG
GGAAGCAGAGGTTGCAGTGAAGTGAATAGTGATAGTGCCACTGCAATTACAGCTGGGTGAC
ATAGAGAGACTCCATCTCAAAAAA

FIGURE 55

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45415

<subunit 1 of 1, 215 aa, 1 stop

<MW: 24326, pI: 6.32, NX(S/T): 4

MHRDAWLPRPAFSLTGLSLFFSLVPPGRSMEVTVPATLNVLNGSDARLPCTFNSCYTVNHKQ
FSLNWTYQECNNCSEEMFLQFRMKIINLKLERFQDRVEFSGNPSKYDVSVMRLRNVQPEDEGI
YNCYIMNPPDRHRGHGKIHLQVLMEEPPEPDRDSTVAVIVGASVGGFLAVVILVLMVVKCVRRK
KEQKLSTDDLKTEEEGKTDGEGNPDDGAK

Important features:

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 161-179

Immunoglobulin-like fold:

amino acids 83-127

N-glycosylation sites.

amino acids 42-45, 66-69 and 74-77

GTTGTATATGTCCTGAAGTACATCCGTGCATTTTTTTTTAGCATCCAACCATCCTCCCTTGTA
 GTTCTCGCCCCCTCAAATCACCTTCTCCCTTAGCCACCCNACTAACATCTCAGTCTCTGAA
 AATGCACAGAGATGCTCGGCTACCTCGCCCTGCCTTACGCCCTACGGGGCTCAGTCTCTTTT
 TCTCTTTGGTGCCACCAAGGACGGAGCATGGAGGTCCACGTACTCTGNCCACCCTCAACGTCC
 TCAATGGCTCTGACGCCCGCCTGCCTTCAACTCCTGCTACACAGTGAACCACAAAC
 AGTTCTCCCTGAACTGGACTTACCAGGAGTGCAACAACCTGCTCTGAGGAGATGTTCTCCAG
 TTCGCATGAAGATCATTAACCTGAGCTGGAGCGGTTTCAAGACCGCGTGGAGTTCTCAGG
 GAACCCACGCAAGTACGATGTGTCGGTGATGCTGAGAAACGTGACGCCGAGGAGTGGGGGA
 TTTCAACTGCTACATCATGAACCCCCC

FIGURE 57

TCACGGGGCTCATCTCTTTTTCTCTTTGGTGCCACCAGGACGGAGCATGGAGGTNCACATA
CCTGCCACCCCTCAACGTCCTCAATGGCTTTGACGCGCGCCTGCCCTGCACCTTCAACTCCNG
CTACACAGTGAACCACAAACAGTTCTCCCTGAACTGGATTACCAGGAGTGCAACAACCTGGC
TCTGAGGAGATGTTCTCCAGTTCCCGCATGGAAGATCATTAAACCTGAAAGCTGGAAGCGG
TTTTCAAGAACCGCGTGGAAGTTTCTCAGGGAACCCAGCAAGTACGATGTGTGGTGATGC
TGAGAAACGTGCAGCCGGAGGATGAGGGGATTACAACCTGCTACATCATGAACCCCC

FIGURE 58

TCGCGCGACCGTCGTACACCA**Atg**GGCCCTCCACCTCCGCCCCACCGTGTGGGGCTGCTCCCGGATGGCCCTCCTGT
 TCCTCTCTGCTGCTGCTAATGCTGCTCGCGGACCCAGCGCTCCCGGCCGGAAGTCAACCCCCAGTGGTGTGCTGCTCC
 CTGGTGATTGGGGTAACCAACTGGAAGCCAAAGCTGGACAAGCCGACAGTGGTGCACTACTCTCTCTCCAAAGAGA
 CCGAAAGCTACTTCACAATCTGGCTGAACCTGGAACTGCTGCTGCTGCTCATATTGACTGCTGGATTGACAATA
 TCAGGCTGGTTTACAACAAAACATCCAGGGCCACCCAGTTTCTGATGGTGGGATGTACGTGTCCCTGGCTTTG
 GGAAGACCTTCTCACTGGAGTTTCTGGACCCAGCAAAAGCAGCGTGGGTTCCTATTTCACACCATGGTGAGAGA
 GCCTTGTGGGCTGGGGCTACACACGGGGTGAGGATGTCCGAGGGCTCCCTATGACTGGCGCCGAGCCCCAAATG
 AAAACGGGCCCTACTTCTGGCCCTCCGCGAGATGATCGAGGAGATGTACAGCTGTATGGGGGCCCCCTGGTGC
 TGGTTGCCCAAGTATGGGCAACATGTACACGCTCTACTTTCTGAGCGGACGCCGAGGGCTGGAAAGACAAGT
 ATATCTGGGCCCTTGTGTCACTGGGTGGCCCTGGGGGGGCTGGGCCAAGACCTCGCGGTCTGGCTTCAGGAG
 ACAACAAACCGGATCCCACTGTCAGGCCCCCTGAAGATCGGGAGCAGCAGCGGTGAGCTGCTCCACAGCTGGC
 TGTCTGCCCTACAACATACATGGTCACCTGAGAAGGTGTTCTGTGAGACACCCACATCACTACACACTGCGGG
 ACTACCGCAAGTTCTCCAGGACATCGGCTTTGAAGATGGCTGGCTCATGCGGCAGGACACAGAAGGGCTGGTGG
 AAGCCAGATGCCACCTGGCGTGACGCTGCATGCTCTCATGTACTGGCTCCCAACACAGACTCCTTCTACT
 ATGAGAGCTTCCCTGACCGTGACCTAAAATCTGCTTTGGTGACGGCGATGGTACTGTGAACCTGAAGAGTGCC
 TGCAGTGCCAGGCTGGCAGAGCCGCGAGGAGCACCAAGTGTGTGTCGAGGAGCTCCAGGACAGGACACATCG
 AGATGCTGGCCAAAGCCACACCTCGGCCATCTGAAACGTGTGCTCCTTGGGCCCT**Tga**CTCCTGTGCCACAGGA
 CTCTGTGTGGCTCGGCCGTGGACCTGCTGTTGGCTCTGGGGCTGTCTAGGCCACCGCGTTTGGCAAGTTTGTGA
 CTCACCATTCAGAGCCCCGAGTCTTGGACTGTGAAGCATCTGCCATGGGGAAGTGTGTTGTTATCCTTCTCT
 GTGGCAGTGAAGAAGGAAGAAATGAGAGCTAGACTCAAGGACACTGGATGCAAGATGCTGCTGATGGTGGGA
 ACTGTGTGACCTTAGGACTGGCTCCACAGGGTGGACTGGCTGGGCCCTGGTCCAGTCCCTGCTGGGGCCATG
 TGTCCCCCTATTCTGTGGGCTTTTCATACTGGCTACTGGGCCCTGGCCCCGACGCTTCTATGAGGGATGTT
 ACTGGGCTGTGTCCTGTACCCAGAGGTCACAGGATCGGCTCCTGGCCCCCTGGGTGACCTTCCACACACCA
 GCCACAGATAGGCCCTGCACCTGGTCATGGGTAGCTAGAGCTGCTGGCTTCCCTGTGGCTTAGCTGTGGCCAGCC
 TGACTGGCTTCTGGCGAGCCTAGTAGCTCCTGCGAGGAGGGGACAGTTGTTGCGTCTCTCTGTGGTCCAGGC
 CCTGGGACATCTCACTCCTCACTCCTCCCTTACCACAGGAGCACTCAAGCTCTGGATTGGGCGACAGATGTG
 CCCCAGTCCCGAGGCTGTGTTCCAGGGGCCCTGATTCTCTCGGATGTCTATGGCCCCAGGACTGAAGCTGC
 CTCCCTTCACTCGGACTGTGGTTCCAAGGATGAGAGCAGGGGTTGGAGCCTTGGGCAACCTATGGA
 GAAAGGAATCCAAGGAAGCAGCCAAAGCTGCTCGAGCTTCCCTGAGCTGCACCTCTTGTCTAACCCACCATCA
 CACTGCCACCTGCCCCTAGGGTCTCACTAGTACCAAGTGGGTGAGCAGGGCTGAGGATGGGGCTCTTATCCAC
 CCTGGCCAGCACCCAGCTTAGTGTGGGACTAGCCAGAACTTGAATGGGACCTGAGAGAGCCAGGGGTCCCC
 TGAGGCCCCCCCTAGGGCTTTTGTGTGTCGCCAGGGTGCTCCATGGATCTCCCTGTGGCAGCAGGATGGAGAGT
 CAGGGCTGCCCTTATGTCAGTGGCTTAAGTGGGTGACTGGCCACAGGCCAGAAAAGGGTACAGCCTCTAGGT
 GGGGTTCCCAAAGACGCTTCAGGCTGGACTGAGCTGCTCTCCACAGGGTTTCTGTGCACTGGATTTCCTGTG
 TTGCATACATGCCTGGCATCTGTCTCCCTTGTCTGTGAGTGCCCCACATGGGGCTCTGAGCAGGCTGTATCTG
 GATTCTGGCAATAAAGTACTCTGGATGCTGTAAAAA

FIGURE 59

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44189
><subunit 1 of 1, 412 aa, 1 stop
><MW: 46658, pI: 6.65, NX(S/T): 4
MGLHLRPYRVGLLPDGLLFLLLLMLLADPALPAGRHPFVVLVPGDLGNQLEAKLDKPTVVH
YLCSKKTESYFTIWLNLLELLLPVIIDCWIDNIRLVYNKTSRATQFPDGVDRVPGFGKTFSL
EFLDPSKSSVGSYFHTMVESLVGWGYTRGEDVRGAPYDWRRAPNENGPYFLALREMIEMYQ
LYGGFVVLVAHSMGNMYTLYFLQRQPQAWKDKYIRAFVSLGAPWGGVAKTLRVLASGDNMRI
PVIGPLKIREQQRSVSTSWLLPYNYTWSPEKVFVQTPTINYTLRDYRKFFQDIGFEDGWL
RQDTEGLVEATMPPGVQLHCLYGTGVPTPDSFYYESFPDRDPKICFGDGTVNLKSALQCQ
AWQSRQEHQVLLQELPGSEHIEMLANATTLAYLKRVLGPF
```

Important features:

Signal peptide:

amino acids 1-28

Potential lipid substrate binding site:

amino acids 147-164

N-glycosylation sites.

amino acids 99-102, 273-276, 289-292 and 398-401

Lipases, serine proteins

amino acids 189-201

Beta-transducin family Trp-Asp repeat

amino acids 353-365

FIGURE 60

CGGACGCGTGGGCGGACGCGTGGGGCGGCGGCAGCGGCGGCGACGCGGACATGGAGAGCGGG
GCCTACGGCGCGGCCAAGGCGGGCGGCTCCTTCGACCTGCGGCGCTTCCTGACGACGCGCA
GGTGGTGGCGCGCGCCGTGTGCTTGGTCTTCGCCTTGATCGTGTCTCTCTGCATCTATGGTG
AGGGCTACAGCAATGCCACGAGTCTAAGCAGATGTACTGCGTGTTC AACCGCAACGAGGAT
GCCTGCGCTATGGCAGTGCCATCGGGGTGCTGGCCTTCCTGGCCTCGGCCTTCTTCTTGGT
GGTCGACGCGTATTTCCCCAGATCAGCAACGCCACTGACCGCAAGTACCTGGTCA TTGGTG
ACCTGCTCTTCTCAGCTCTCTGGACCTTCCTGTGGTTTGTGGTTTCTGCTTCCTCACCAAC
CAGTGGGCAGTCACCAACCCGAAGGACGTGCTGGTGGGGGCCGACTCTGTGAGGGCAGCCAT
CACCTTCAGCTTCTTTCCATCTTCTCCTGGGGTGTGCTGGCCTCCCTGGCCTACCAGCGCT
ACAAGGCTGGCGTGGACGACTTCATCCAGAATTACGTTGACCCCACTCCGACCCCAACACT
GCCTACGCCTCCTACCCAGGTGCATCTGTGGCAACTACCAACAGCCACCCCTTCAACCAGAA
CGCGGAGACCACCGAGGGCTACCAGCCGCCCCCTGTGTA~~CTG~~AGTGGCGGTTAGCGTGGGAA
GGGGGACAGAGAGGGCCCTCCCCCTGCCCCTGGACTTTCCCATCAGCCTCCTGGAAC TGCCA
GCCCTCTCTTTACCTGTTCATCCTGTGACAGTGACACACAGCTAAGGAGCCTCATAGCC
TGCGGGGGCTGGCAGAGCCACCCCCAAGTGCCCTGTGCCAGAGGGCTTCAGTCAGCCGCT
CACTCCTCCAGGGCACTTTTAGGAAAGGGTTTTAGCTAGTGTTTTCTCTCGCTTTTAATGA
CCTCAGCCCCGCTGCAGTGGCTAGAAGCCAGCAGGTGCCATGTGCTACTGACAAGTGCCT
CAGCTTCCCCCGGCCGGGTCAGGCCGTGGGAGCCGCTATTATCTGCGTTCTCTGCCAAAG
ACTCGTGGGGGCCATCACACCTGCCCTGTGCAGCGGAGCCGACCAAGGCTCTTGTGTCTTCA
CTCAGGTTTGCTTCCCCCTGTGCCCACTGCTGTATGATCTGGGGGCCACCAACCTGTGCCGCT
GGCCTCTGGGCTGCCCTCCCGTGGTGTGAGGGCGGGGCTGGTGTCTATGGCACTTCTCTCTTG
CTCCACCCCTGGCAGCAGGAAGGGCTTTGCCTGACAACACCCAGCTTTATGTAATATTC
TGCAGTTGTTACTTAGGAAGCCTGGGGAGGGCAGGGGTGCCCCATGGCTCCAGACTCTGTG
TGTGCCGAGTGATTATAAAATCGTGGGGAGATGCCCGGCTGGGATGCTGTTTGGAGACG
GAATAAATGTTTTCTCATTCAAAG

FIGURE 61

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48304
<subunit 1 of 1, 224 aa, 1 stop
<MW: 24810, pI: 4.75, NX(S/T): 1
MESGAYGAAKAGGSFDLRRFLTQPQVVARAVCLVFALIVFSCIYGE GYSNAHESKQMYCVFN
RNEDACRYGSAIGVLAF LASAFFLVVDAYFPQISNATDRKYL VIGDLLFSALWTF LWFVGFC
FLTINQWAVINPKDVLVGADSVRAAITFSFFSIFSWGVLASLAYQRYKAGVDDFIQNYVDPTF
DPNTAYASYPGASVDNYQQPPFTQNAETTEGYQPPFVY

Important features:

Type II Transmembrane domain:

amino acids 1-45

Other transmembrane domains:

amino acids 74-90, 108-126 and 145-161

N-glycosylation site.

amino acids 97-100

FIGURE 62

GAGCCACCTACCTGCTCCGAGGCCAGGCCCTGCAGGGCCTCATCGGCCAGAGGGTGATCAGTGAGCAGAAGGATG
 CCCGTGGCCAGGGCCCCCAGGTGGCTTGGCGGCAGGGGGACGGAGGTGATGGCAGGGAAGCGAGCCAGAGGGG
 ATGTTCAAGGCTGTGAGGACTCCAAGAGAAAAGCCCGGGCTACTCCGCTGGTGCCCTGTGTTGTGCTGCTG
 GCCTGCTGTGCTGGCTTCCGCGGGGGTGCTACTCTGGTAATTCTTAGGTACAAAGCGAGGTGATGGTCAGC
 CAGGTGTACTCAGGCAGTCTGGGTGTACTCAATCGCCACTTCTCCAGGATCTTACCCGCGCGGGAATCTAGTGCC
 TTCGCAGTGAACCCGCCAAGGCCAGAAAGTGTCTAAGAGAGTCTATCACCAGCACCCGCTTGGGAATCTACTAC
 AACTCCAGTCCCGTCTATTCTTTGGGAGGGACCCCTCACCTGCTTCTTCTGGTTCAATCTCCAAATCCCCGAG
 CACCGCGGCTGTAGTCTGAGCCCCAGGTGGTGCAGGCACCTGCTGGTGGAGAGGAGTGTCTGTCCAGTCAACAGC
 TCGGCTGCGCTCCCTCAGGGGCCAGTACGAAGTGGACCCCGAGGGCCTAGTGTCTTGGAAAGCCAGTGTGAAA
 GACATAGTGTCAATGAATTCACGCTGGGTGTGTACCGCTACAGCTACGTGGGCGAGGGCAGGTCTCTCCGGT
 AAGGGCTGTACCACTGTGCCTTCAGCTGGCTTGTGGCACTGCAGGGCCCCAAGGACCTCATGTCTCAAACTCCG
 CTGCAGTGGAGCTGGCAGAGTGCCTGGGACCGACTGGCCATGTATGACGTGGCCGGGCCCCGGAGAAGAGGCTC
 ATCACTCGGTGTACGGCTGCAGCGCCAGGAGCCCGTGGTGGAGGTTCTGGCTCGGGGCCATCATGGCGGT
 GTCTGGAAGAAGGGCTTGCACAGCTACTACGACCCCTCTGCTCTCTCCGTGCAGCCGCTGGTCTTCCAGCCCTGT
 GAAGTGAACCTGACGCTGGACAAAGGCTCGACTCCAGGGCGTCTCTGCAGCACCCGCTACTTCCCCAGCTACTAC
 TCGCCCCAAACCCACTGTCTCTGGCACTCAGGTGCCCTCTCTGGACTACGGCTTGGCCCTCTGGTTTGATGCC
 TATGCACTGAGGAGGCAGAAATGATGTTTCCGCTGCACCAGGGCCAGTGGAGCATCCAGAACAGGAGGCTGTGT
 GGTCTGCGATCCTTGACGCCCTACGCCGAGAGGATCCCGTGGTGGCCACGGCCGGATCACCATCAACTTCACC
 TCCAGATCTCCCTCACCGGGCCCGGTGTGCGGGTGCACTATGGCTGTACAACAGTCCGAGCCCCCTGCCCTGGA
 GAGTTCCTCTGTTCTGTGAATGGACTGTGTCCCTGCTGTGTGAGGGTCAAGGACTGCCCAACGGCGCTGGAT
 GAGAGAACTGGCTTTCAGAGGCCAATTCCAGTGCAGAGGAGCAGCACATGCTATCTCATGTGCCAAGGTCTGT
 GATGGGCGCCTGATTGTCTCAACGGCAGCGATGAAGAGCAGTGCAGGAAGGGGTGCCATGTGGGACATCAAC
 TTCAGTGTGAGGACCGGAGTGGTGAAGAAGCCCAACCCGAGTGTGATGGGCGGCCGCACTGCAGGAGCGGC
 TCGGATGAGGAGCACTGTGACTGTGGCTCCAGGGCCCCCTCAGCGCATTGTTGTGTGAGCTGTGTCTCCGAG
 GGTGAGTGGCCATGGCAGGCCAGCTCCAGGTTCCGGGTGCACACATCTGTGGGGGGGCCCTCATGCTGACCGC
 TGGGTGATAACAGCTGCCCACTGCTTCCAGGAGGACAGCATGGCCCTCCACGGTGTGTGGACCGGTGTTCTGGGC
 AAGGTGTGGCAGAACTCGCGCTGGCCTGGAGAGGTGTCTTCAAGGTGAGCCGCTGTCTCTGCACCCGTAACAC
 GAAGAGGACAGCCATGACTACGAGCTGGCGTGTCTGAGCTCGACCAACCGGTGTGTGCGCTCGGCCCGGTGCGC
 CCGTCTGCTGCCCGCGCTCCCACTTCTTCAGAGCCCGGCTGCACTGTGGATTACGGGCTGGGCGCGCTTG
 CGGAGGGCGGCCCCATCAGCAACGCTCTGCAGAAAGTGGATGTGAGTTGATCCACAGGACCTGTGCAGCGAG
 GCCTATCGTACCAAGTGAACGACAGCATGTGTGTGCGGCTACCGCAAGGGCAGAAAGGATGCTGTCAAGGT
 GACTCAGGTGGTCCGCTGGTGTGCAAGGCACTCAGTGGCCGCTGGTTCTTGGCGGGGTGGTCAAGTGGGGCTG
 GGCTGTGGCCCGCTTAACCTACTTCCGCGCTACACCCGATCACAGGTGTGATCAGCTGGATCCAGCAAGTGGT
 ACCTGAGGAATCGCCCCCTGCAGAAAGGGGCCACCTCTCGGACTCAGAGAGCCAGGGCACTGCCAAGCAGG
 GGGACAGTATTTCTGGCGGGGGTGGGGAGAGAGCAGGCCCTGTGGTGGCAGGAGGTGGCATCTGTGTCTGTC
 CTGATGTCTCTGCTCAGTGTGCGAGGAGTGGAGAAGTGCACAGCTGGAGGGTCAAGAGCTCTTGCATCAGAG
 CAGGCCACACCCAGCCCTTCTGCTCCCAATCTCTCTCTCCGTCCCTTCTCTCACTGCTCTCTATGTACAG
 GCAGTCCCTCAGCAGCAAGATGCTGGTGTCTCATCTCCGAGGAGTGTCTAGTGTGGCGCCCATCTGTATCAGG
 CTGTTTGGGACGCTTGCCTCCAGAGAGCAGATTCCAGCTTCCGAAGCCCCCTGGTCTAATCTTGGGATCTGGGAAT
 GGAAGGTGCTCCCATCCGAGGGGCCCTCAGAGCTCCGAGGTGGGCTGTGCTGCCATCTGAAGCCTTAAGCCAA
 AAGGTGGGAACTCTGACTCCAGGGTCTTGGCCCCACCCCTGCTGCCACCTGGGCCCCACAGCCGACGCTC
 CACTGGAGGTGAGCTCAGTGCCTTTGGAATAAGCTGCTGTATCAAAAAAAAAAAAAAAAAAAAA

FIGURE 63

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49152
><subunit 1 of 1, 802 aa, 1 stop
><MW: 88846, pI: 6.41, NX(S/T): 7
MPVAEAPQVAGGQDGGDGEEAEPEGMFKACEDSKRKARGYLRVLVPLFVLLALLVLASAGVL
LWYFLGYKAEVMVSQVYSGSLRVLNRHFSQDLTRRESSAFRSETAKAQKMLKELITSTRLGT
YYNSSSVYSFGEGLTCFFWFILQIPEHRRMLLSPEVVQALLVEBLLSTVNSSAAVPYRAEY
EVDPEGLVILEASVKDIAALNSTLGCYRYSYVGQGVRLRLKGPDLHASSCLWHLQGPDLML
KLRLEWTLAECDRLAMYPDAGPLEKRLITSVYGCSRQEPVVEVLASGAIMAVVWKKGLHSY
YDPFVLSVQPVVFQACEVNLTLDNRLDSQGVLSPTYFSPYSYSPQTHCSWHLTVPSLDYGLAL
WFDAYALRRQKYDLPCQTQGGWTIQNRRLCGLRILQPYAERIPVVATAGITINFTSQISLTGP
GVRVHYGLYNQSDPCPGEFLCSVNGLCVPACDGVKDCPNGLDERNCVCRATFQCKEDSTCIS
LPKVCDCGQPDCLNGSDEEQCQEGVPCGTFTFQCEDRSCVKPNPQCDGRPDGRDGSDEEHCD
CGLQGPSSRIVGGAVSSEGEWPWQASLQVRGRHICGGALIA DRWVITA AHCFQEDSMAS TVL
WTVFLGKQVQNSRWPGEVSFKVSRLLLHPYHEEDSHDYDVALQLDHPVVRSAAVRPVCLPA
RSHFFEPGLHCWITGWGALREGGPISNALQKVDVQLIPQDLCSEAYRYQVTPRMLCAGYRKG
KKDACQGDSSGGPLVCKALSGRWFLAGLVSWGLGCGRPNYFGVYTRITGVISWIIQQVVT
```

Important features:

Type II transmembrane domain:

amino acids 46-67

Serine proteases, trypsin family, histidine active site.

amino acids 604-609

N-glycosylation sites.

amino acids 127-130, 175-178, 207-210, 329-332, 424-427, 444-447
and 509-512

Kringle domains.

amino acids 746-758 and 592-609

Homologous region to Kallikrein Light Chain:

amino acids 568-779

Homologous region to Low-density lipoprotein receptor:

amino acids 451-567

FIGURE 64

GCACCCAGGGCCAGTGGACGATCCAGAACAGGAGGCTGTGTGGCTTGCGCATCCTGCAGCCC
TACGCCGAGAGGATCCCCGTGGTGGCCACGGCCGGGATCACCATCAACTTCACCTCCCAGAT
CTCCCTACCGGGCCCGGTGTGCGGGTGCACTATGGCTTGTAACAACAGTCGGACCCCTGCC
CTGGAGAGTTCTCTGTTCTGTGAATGGACTCTGTGTCCCTGCCTGTGATGGGGTCAAGGAC
TGCCCCAACGGCTGGATGAGAGAACTGCGTTTGACAGGCCACATTCCAGTGCAAAGAGGA
CAGCACATGCATCTCACTGCCAAGGTCTGTGATGGGCAGCCTGATTGTCTCAACGGCAGCG
ATGAAGAGCAGTGCCAGGAAGGGGTGCCATGTGGGACATTACCTTCCAGTGTGAGGACCGG
AGCTGCGTGAAGAAGCCCAACCCGCAGTGTGATGGGCGGCCGACTGCAGGGACGGCTCGGA
TGAGGAGCACTGTGACTGTGGCTCCAGGGCCCCCTCCAGCCGCATTGTTGGTGGAGCTGTGT
CCTCCGAGGGTGAGTGGCCATGGCAGGCCAGCCTCCAGGTTCCGGGTTCGACACATCTGTGGG
GGGGCCCTCATCGCTGACCGCTGGGTGATAACAGCTGCCCACTGCTTCCAGGAGGACAGCAT
GGCCTCCACGGTGCTGTGGACCGTGTTCTTGGGCAAGGTGTGGCAGAAGTTCGCGCTGGCCTG
GAGAGGTGTCTTCAAGGTGAGCCGCTGCTCCTGCACCCGTACCAACGAAGAGGACAGCCAT
GACTACGACGTGGCGCTGCTGCAGCTCGACCACCCGGTGGTGCCTCGGCCGCGCTGCGCCC
CGTCTGCCCTGCCCGCGCTCCCACTTCTTCGAGCCCGCCTGCACTGCTGGATTACGGGCT
GGGGCGCCTTGCGCGAGGGCGGCCCATCAGCAACGCTCTGCAGAAAGTGGATGTGCAGTTG
ATCCCAAGGACCTGTGCAGCGAGGCCTATCGCTACCAGGTGACGCCACGCATGCTGTGTGC
CGGTAACCGCAAGGGCAAGAAGGATGCCTGTGAGGGTGACTCAGGTGGTCCGCTGGTGTGCA
AGGCACTCAGTGGCCGCTGGTTCTTGGCGGGGCTGGTCAGTGGGGCCTGGGCTGTGGCCGG
CCTAACTACTTCGGCGTCTACACCCGCATCACAGGTGTGATCAGTGGATCCAGCAAGTGGT
GACCTGAGGAAGTGGCCCCCTGCAAAGCAGGGCCACCTCCTGGACTCAGAGAGCCAGGGC
AACTGCCAAGCAGGGGGACAAGTAT

Figure 1. Schematic representation of the experimental design. The subjects were divided into two groups: the control group (CG) and the experimental group (EG). The CG was divided into two subgroups: the control group (CG) and the control group (CG). The EG was divided into two subgroups: the experimental group (EG) and the experimental group (EG). The subjects were divided into two groups: the control group (CG) and the experimental group (EG). The CG was divided into two subgroups: the control group (CG) and the control group (CG). The EG was divided into two subgroups: the experimental group (EG) and the experimental group (EG).

[illegible]

FIGURE 66

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49646

><subunit 1 of 1, 354 aa, 1 stop

><MW: 39362, pI: 8.35, NX(S/T): 2

MSNSVPLLCFWSLCYCFAAGSPVPFGPEGRLEDKLHKPKATQTEVKPSVRFNLRTSKDPEHE
GCYLSVGHSQPLEDCSFNMTAKTFFIIHGWTMSGIFENWLHKLVSALHTREKDANVVVDWL
PLAHQLYTDVNNTRVVGHSIARMLDWLQEKDDFSLGNVHLIGYSLGAHVAGYAGNFVKGTV
GRITGLDPAGPMFEGADIIHKRLSPDDADFVDVLHTYTRSFGLSIGIQMPVGHIDIYPNGGDF
QPGCGLNDVLGSIAYGTITEVVKCEHERAVHLFVDSL VNQDKPSFAFQCTDSNRFKKGICLS
CRKNRCNSIGYNAKMRNKRNSKMYLKTRAGMPFRGNLQSLECP

Important features:

Signal peptide:

amino acids 1-16

Lipases, serine active site.

amino acids 163-172

N-glycosylation sites.

amino acids 80-83 and 136-139

FIGURE 67

CGGACGCGTGGGCGGACGCGTGGGCTGGGCAAGGGCGGGGGCGCGGGCGAGGCCACCTCTTCCCCCCCCCG
 TTCCTTGTCCGGCTCCGCTGGCTGGAGCGCTGGAGGAGTGGAGCAGCACCCGGCGCGCTTGGGGGTGACAGT
 CGGCCAAGTGTGGGCGGAGAGGAAAGTGGTCTCAAACCCGGCAGGTGGCGACACGGCCAGACAGGGGCGCTCG
 CTGCTTGGCGGGCGGGCTGTAGGCGAGGGGCGCGCCCAAGTCCCGAGACCCGGGGCTTCAGGAGCGCGCCCGGGAG
 AGAAGAGTGGCGGGCGGACGAGGAAACAACTCCAAAGTTGGCGAAAGGACCCGCCCTACTCCCGGGCTGGCG
 CCGCTTCCCGCCCCCAGCCCTGGCATCCAGAGTACGGGTTCGAGCCCGGGCCATGGAGCCCGCTGGGAGGGCGG
 CACACGGAGCTGGGCGCCCGGGGCTCCGCGCGACCCCATCGGGTAGACACAGAAGCTTCGGGACCCCTTCG
 GCACCTTCTGGACAGCCGAGTGTCTGTGGCCACCTCCTCTCTCTCTCTCTGGAGGGCTCTGGCCCATCCAG
 ACCGATATTATTTTCCAAATCATGCTTGTGAGGACCCCCCAGCAGTGTCTTTAGAAGTGCAGGGCACCTTACAGA
 GGGCCCTGGTCCGGGACAGCCCACTCCCTTGCCTGCACTGCACCTGGCTCATCTGGGCGACGAAGAACAGACT
 TCAACATCAGTTCAGGAGCTTACACCTGGCCCTTGGCTCAGAGGGCTTAACTCCCTCCCTTCCCTCAGCGCTG
 TGAATCTCCCTGTGTGAGGCACTCCAGCCCTCTGCAGCTGCCCGGGGCAAGCTCACCATCACTTACAGCTATG
 CTGGGGCCAGAGCACCCATGGGCGAGGGCTTCTGCTCTCTACAGCCAAAGATTGGCTGATGTGCTCGCAGGAAG
 AGTTTCAGTGCTGAACACCCGCTGTGATCTGTCTGTCCAGCGCTGTGATGGGTTGATGCTGTGGCGATGGCT
 CTGATGACGAGGTTGCAGCTCAGACCCCTTCCCTGGCCTACCCCAAGACCGCTGCCCTCCCTGCTTGCAGT
 TCACCTTGGAGGACTTCTATGGGGTCTTCTCTCTCTCTGGATATACACACTAGCCTCAGTCTCCACCCCCAGT
 CTTGCCATTTGGCTGTGACCCCATGATGGCGGGCGCTGGCGGTGCGCTCACAGCCCTGGACTTGGGCTTTG
 GAGATGCAGTGCATGTGTATGACGGCCCTGGGCCCTTGAGAGCTCCCGACTACTCGTGTGTCTACCCCACTTCA
 GCAATGCAGAGGCTGTCACTGTGGAGACACTGTCTGGCCAGGCTGTGTGTCTTACACACAGTTGCTTGGAGCA
 ATGGTCTGGGCTTCAATGCCACTACCATGTGCGGGGCTATTGCTTGCTTGGGACAGACCTGTGGCTTAGGCT
 CTGGCTGGGAGCTGGCGAAGGCTAGGTGAGCGCTGTACAGTGAAGGACAGCGCTGTGACGGCTCATGGGACT
 GTGCTGACGGCAGATGAGGAGGACTGCCAGGTGCCACCTGGACACTTCCCTGTGGGGCTGCTGGCACT
 CTGGTGCCACAGCTGCTACCTGCTGCTGACCGCTGCAACTACAGACTTCTGTGTGATGGAGCAGATGAGA
 GAGCTGTGCGCATTTGCCAGCTGGCAATTTCCGATGCCGGGACGAGAAGTGGGTGTATGAGACGTGGGTGTGG
 ATGGGACGCCAGACTGTGCGGACGGCAGTGATGAGTGGGACTGCTCTATGTTCTGCCCGCAGGTCATTACAG
 CTGCACTATTGGCAGCCTAGTGTGCGGCTGCTCCTGGTCACTGCCCTGGGCTGCACCTGCAAGCTCTATGCCA
 TTGCAACCCAGGAGTACAGCATTTTGGCCCCCTCTCCCGGATGGAGGCTGAGATTGTGACAGCAGGACACCC
 CTTCTACGGGCACTCATTTGCCAGGGTGCCATCCCACTGTAGAAGACTTTCCTACAGAGAACTCTAATGATA
 ACTCAGTGCTGGGCAACCTGCGTTCTCTGCTACAGATCTTACGCCAGGATATGACTCCAGGAGGTGGCCAGGTG
 CCGCGCTGTGTCAGCGGGGCGGCTTGTATGCGACGCTGGTACGCGCTCTCCGCGCTGGGGCTTGGCTCCCTCGAA
 CCAACACCCCGGCTCGGGCTCTGAGGCCAGATCCCAAGTCAACCTTCTGTGCTCCCTTGAAGCCCTGAGT
 GTGGCAGAGTCCAGCCCTGAGGGCGGGCAGTGGGTGGGCAAGATGGGGAGCAGGACCCCACTGCCATCA
 AGGCTCCCTCCATCTGTAGCAGCTTCCAGCCCCACTACTGTCCCTGAAGCCCGAGGCGACTGCCCTCAC
 TGGCCCTAGAGCCATCATATTGTCTGGAGTGGTGCAGGCCCCGCGAGGCTGGTTCGCCAGCTGGGGCCCC
 CAGGACCAACCCGAGCCCCCTGGACCCCAACAGCAGTCTGGCCCTGGAAGATGAGGACGATGTGCTACTGG
 TGCCACTGGCTGAGCCCGGGGTGTGGGTAGCTGAGGCAGAGGATGAGCCACTGCTTACTTGAAGGAGCTGGGG
 CTCCTACTGAGGCTCTCCCTGGGGGCTCTACTCATAGTGGCAACACTTTTGAAGGTGGTGTGAGCTCCCTCC
 ACCACTTCTTCCCTGCTTCCCTGGATTTCAGGGACTTGGTGGGCTTCTTTTAAAGTGAAGTTCTTAGAGGATCAT
 TAACTGTCCCTCAGGACAGGAGGGCTCACAGAGTCTCTCTGTGACGTGGCCATGGCCAGACACCCAGTCCCT
 TCACACCACTGCTCCCAACCCCAACCAATTGGGTGGCTGTCTTTTAAAGTGAAGTTCTTAGAGGATCAT
 GGTCTGGACCTCCATCTTGCCTAACCTCTACCCAAAGTGGCTTAAAGCAGGGAATGCCAATTAACATGAGA
 CCTCCAGCCCCCAAGGGAGGATTTGGGAGAACCTGAGGTTTTGCCATCCCAATCCCTCTCAGGGGCTGG
 CTCACAAAGAGTGCACAAATGCTTCTATTCATAGCTACGGCATTGCTCAGTAAGTTGAGGTCAAAATAAA
 GGAATCATACATCTC

CGGACGCGTGGGCGGACGCGTGGGCTGGGCAAGGGCGGGGGCGCGGGCGAGGCCACCTCTTCCCCCCCCCG

FIGURE 68

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49631

<subunit 1 of 1, 713 aa, 1 stop

<MW: 76193, pI: 5.42, NX(S/T): 4

MLLATLLLLLLGGALAHDPRIIFPNHACEDPPAVLLEVQGTLQRPLVRDSRTSPANCTWLIL
GSKEQTVTIRFQKLHLACGSERLTLRSPLQPLISLCEAPPSPLQLPGGNVTITYSYAGARAP
MGQGFLLSYSQDWLMCLQEEFQCLNHRCVSAVQRCGDVDACGDGSDEAGCSSDFFPGLTPRP
VPSLPCNVLTLEDIFYGVFSSPGYTHLASVSHQPQSCHWLLDPHDGRR LAVRFTALDLGFDAVH
VYDGPGPPESSRLRLSLTHFSNGKAVTVETLSGQAVVSYHTVAWSNGRGFNATYHVRGYCLP
WDRPCGLGSGLGAGEGLGERCYSEAQRCDGSDCADGTDEEDCPGCPPGHFPCGAAGTSGAT
ACYLPADRCNYQTFCADGADERRCRHCQPGNFRRCRDEKCVYETWVCDGQPD CADGSDEWDCS
YVLPRKVITAAVIGSLVCGLLLVIALGCTCKLYAIRTQEYSIFAPLSRMEAEIVQQQAPPSY
GQLIAQGAIPPVEDFPTENPNDNSVLGNLRSLLQILRQDMTPGGGPGARRRQRGRLMRRLVR
RLRRWGLLPRTNTPARASEARSQVTPSAAPLEALDGGTGPAREGGAVGGQDGEQAPPLPIKA
PLPSASTSPAPTTVPEAPGPLPSLPLEPSLLSGVVQALRGRLLP SLGPPGPTRSPPGPHTAV
LALEDEDDVLLVPLAEFGVWVAEAEDEPLLT

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 442-462

LDL-receptor class A (LDLRA) domain proteins

amino acids 411-431, 152-171, 331-350 and 374-393

FIGURE 69

FIGURE 70

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49645
><subunit 1 of 1, 152 aa, 1 stop
><MW: 17170, pI: 9.62, NX(S/T): 1
MDNVQPKIKHRPFCFSVKGHVKMLRLALTVTSMTFFIIAQAPEPYIVITGFEVTVILFFILL
YVLRDLRLMKWLFWPLLDIINSLVTTVFMLIVSVLALIPETTTTLTVGGGVFALVTAVCCLAD
GALIYRKLLFNPSGPYQKKPVHEKKEVL
```

Important features:

Potential type II transmembrane domain:

amino acids 26-42

Other potential transmembrane domain:

amino acids 44-65, 81-101 and 109-129

Leucine zipper pattern

amino acids 78-99 and 85-106

N-myristoylation site.

amino acids 110-115

Ribonucleotide reductase large subunit protein

amino acids 116-127

FIGURE 71

GGGCGAGAAGTAGGGGAGGGCGTGTTCGCCGCGGTGGCGGTTGCTATCGTTTTGCAGAACC
TACTCAGGCAGCCAGNTGAGAAGAGTTGAGGGAAAGTGCTGCTGCTGGGCTGCAGACGCGA
TGGATAACGTGCAGCCGAAAAATAAACATCGCCCCCTTCTGCTTCAGTGTGAAAGGCCACGTG
AAGATGCTGCGGCTGGCACTAACTGNGACATCTATGACCTTTTTTATNATCGCACAAGCCCC
TGAACCATATATTGTTATCACTGGATTTGAAGTCACCGTTATCTTATTTTCATACTTTTAT
ATGTACTCAGACTTGATCGATTAATGAAGTGGTTATTTTGGCCTTTGCTTGATATTATCAAC
TCACTGGTAACAACAGTATTATGCTCATCGTATCTGTGTTGGCACTGATACCAGAAACCAC
AACATTGACAGTTGGTGGAGGGGTGTTTGCACTTGTGACAGCAGTATGCTGINTTGCCGAC

5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 72

CAGCCCCGCGCGCCGCCGAGTCTGCTGAGCCGCGGCTGCCGGAACGGACGGGACCGGCTAGG
CTGGGCGCGCCCCCGGGCCCCCGCGTGGGCATGGGCGCACTGGCCCCGGCGCTGCTGCTGC
CTCTGCTGGCCAGTGGCTCTGCGCGCCGCCCGGAGCTGGCCCCCGCGCCCTTCACGCTG
CCCCCTCCGGGTGGCGCGGCCACGAACCGCGTAGTTGCGCCCCACCCGGGACCCGGGACCCC
TGCCGAGCGCCACGCCGACGGCTTGGCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGG
GCGCCGCCAACTTCTTGGCCATGGTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTAC
CTGGAGATGCTGATCGGGACCCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAG
TAACTTTGGCGTGGCAGGAACCCCGCACTCCTACATAGACACGTA¹CTTTGACACAGAGAGGT
CTAGCACATACCGCTCCAAGGGCTTTGACGTACAGTGAAGTACACACAAGGAAGCTGGACG
GGCTTCGTTGGGGAAGACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTGTCAA
CATTGCCACTATTTTTGAATCAGAGAATTCTTTTTGCGCTGGGATTAAATGGAATGGAATAC
TTGGCCTAGCTTATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCC
CTGGTGACACAAGCAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCC
CGTTGCTGGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGTGGAATTGAACCAAGTTTGT
ATAAAGGAGACATCTGGTATACCCCTATTAAAGGAAGAGTGGTACTACCAGATAGAAATTCTG
AAATTGGAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGC
CATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATCGCGGTGGTGG
AAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCCCAG
CTGGCGTGCTGGACGAATTTCGAAAACACCTTGGTCTTACTTCCCATAAATCTCCATCTACCT
GAGAGACGAGAACTCCAGCAGGTCA²TTCCGTATCACAATCCTGCCTCAGCTTTACATTACGC
CCATGATGGGGCCGGCCTGAATTATGAATGTTACCGATTCCGGCATTTC³CCATCCACAAT
GCGTG⁴GGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCACTTCGACAGAGCCAGAA
GAGGTGGGCTTCGACGCGAGCCCTGTGCAGAAATGCAGGTGCTGCAGTGTCTGAAATTT
CCGGGCC⁵TTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCTCAGTCTTTGAGCGAG
CCCATTTTGTGGATTGTGTCTATGCGCTCATGAGCGTCTGTGGAGCCATCCTCCTTGTCTT
AATCGTCTGTGCTGCTGCCGTTCCGGTGTGAGCGTCGCCCCGTGACCCTGAGGTCTGCA
ATGATGAGTCTCTCTGGTCAGACATCGCTGGAATTGAATAGCCAGGCCTGACCTCAAGCAA
CCATGA⁶ACTCAGCTATTAAGAAAAATCACATTTCCAGGGCAGCAGCCGGGATCGATGGTGGCG
CTTTCTCCTGTGCCACCCGCTCTTCAATCTCTGTTCTGCTCCAGATGCCTTCTAGATTAC
TGTCTTTTGATTCTTGATTTTCAAGCTTTCAAATCCTCCCTACTTCCAAGAAAAATAATTAA
AAAAAAACTTCATTCTAA

FIGURE 73

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45493

><subunit 1 of 1, 518 aa, 1 stop

><MW: 56180, pI: 5.08, NX(S/T): 2

MGALARALLLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGLAL
ALEPALASPAGAAANFLAMVDNLQGDSDGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAGTPHS
YIDTYFDTERSSTYRSKGFDTVVKYTGGSWTGFVGEDLVTI PKGFNTSFLVNIATIFESENF
FLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVAGSGTNGGS
LVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKAIVDSGTTLLR
LPQKVFDVAVEAVARASLIPEFSDGFWTGSQLACWTNSETPWSYFPKISIIYLRDENSRSFR
ITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFDRAQKRVGFAASPCA
EIAGAAVSEISGPFSTEDVASNCVPAQSLSEPI LWIVSYALMSVCGAILLVLLVLLLLPFRCL
QRRPRDPEVVNDESSLVRHRWK

Important features:

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 466-494

N-glycosylation sites.

amino acids 170-173 and 366-369

Leucine zipper pattern.

amino acids 10-31 and 197-118

Eukaryotic and viral aspartyl proteases

amino acids 109-118, 252-261 and 298-310

FIGURE 74

CGCCTCCGCCCTTCGGAGGCTGACGCGCCCGGGCGCGTTCAGGCCTGTGCAGGGCGGATCG
GCAGCCGCTTGGCGGCATCCAGGGCGGTGCGGGCCTGGGCGGGAGCCGGGAGGCGCGGCC
GGC**ATG**GAGGCGCTGCTGCTGGGCGCGGGTTGCTGCTGGGCGCTTACGTGCTTGTCTACTA
CAACCTGGTGAAGGCCCGCGTGCGGCGGCATGGGCAACCTGCGGGGCGGCACGGCCGTGG
TCACGGGCGCCAACAGCGGCATCGGAAAGATGACGGCGCTGGAGCTGGGCGCGCGGGGAGCG
CGCGTGGTGTGGCCTGCCGAGCCAGGAGCGCGGGGAGGCGGCTGCCTTCGACCTCCGCCA
GGAGAGTGGGAACAATGAGGTCATCTTCATGGCCTTGGACTTGGCCAGTCTGGCCTCGGTGC
GGCCTTTGCCACTGCCTTTCTGAGCTCTGAGCCACGGTTGGACATCCTCATCCACAATGCC
GGTATCAGTTCCTGTGGCCGACCCGTGAGGCGTTTAACTGCTGCTTCGGGTGAACCATAT
CGGTCCCTTTCTGCTGACACATCTGCTGCTGCCTTGCTGAAGCATGTGCCCTAGCCGCG
TGGTGGTGGTAGCCTCAGCTGCCCACTGTGGGGACGTCTTGACTTCAAACGCTGGACCGC
CCAGTGGTGGGCTGGCGGCAGGAGCTGCGGGCATATGCTGACACTAAGCTGGCTAATGTACT
GTTTGGCCCGGAGCTCGCCAACCAAGCTTGAGGCCACTGCGCTCACCTGCTATGCAGCCACC
CAGGGCCTGTGAACCTCGGAGCTGTTCTGCGCCATGTTCTCGGATGGCTGCGCCCACTTTTG
CGCCCATTTGGCTTGGCTGGTCTCCGGGCACCAAGAGGGGGTGCCAGACACCCCTGTATTG
TGCTCTACAAGAGGGCATCGAGCCCTCAGTGGGAGATATTTTGCCAACCTGCCATGTGGAAG
AGGTGCCCTCCAGCTGCCCCGAGACGACCGGGCAGCCCATCGGCTATGGGAGGCCAGCAAGAGG
CTGGCAGGGCTTGGGCTGGGGAGGATGCTGAACCCGATGAAGACCCCACTGTGAGGACTC
AGAGGCCCCATCTTCTCTAAGCACCCCCCACCCTGAGGAGCCACAGTTTCTCAACCTTACC
CCAGCCCTCAGAGCTCACCAGATTTGTCTAAGATGACGCAACCGAATTCAGGCTAAAGTTGAG
CCTGAGATCCAGCTCTCC**TAA**CCCTCAGGCCAGGATGCTTGCCATGGCACTTCATGGTCTCT
GAAAACTCGGATGTGTGTGAGGCCATGCCCTGGACACTGACGGGTTTGTGATCTTGACCTC
CGTGGTTACTTTCTGGGGCCCCAAGCTGTGCCCTGGACATCTCTTTTCTGGTTGAAGGAAT
AATGGGTGATTATTTCTTCTGAGAGTGACAGTAACCCAGATGGAGAGATAGGGGTATGCT
AGACACTGTGCTTCTCGGAAATTTGGATGTAGTATTTTTCAGGCCCCACCTTATTGATTCTG
ATCAGCTCTGGAGCAGAGGCAGGGAGTTTGCAATGTGATGCATGCCAACATTGAGAATTAG
TGAAGTATCCCTTTGCAACCGCTTAGCTAGGTAGTTAAATTACCCCATGTTAATGAAGCG
GAATTAGGCTCCCGAGCTAAGGGACTCGCCTAGGGTCTCACAGTGAGTAGGAGGAGGGCCTG
GGATCTGAACCAAGGGTCTGAGGCCAGGGCCGACTGCCGTAAAGATGGGTGCTGAGAAGTGA
GTCAGGGCAGGGCAGCTGGTATCGAGGTGCCCATGGGAGTAAGGGGACGCTTCCGGGCGG
ATGCAGGGCTGGGGTCATCTGTATCTGAAGCCCTCGGAATAAAGCGCGTTGACCGCCAAAA
AAAAAAAAAAAAAAAA

FIGURE 75

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48227
<subunit 1 of 1, 377 aa, 1 stop
<MW: 40849, pI: 7.98, NX(S/T): 0
MEALLLGAGLLLGAYVLVYYNLVKAPPCGGMGNLRGRTAVVTGANSIGIGKMTALELARRGAR
VVLACRSQERGEAAAFDLRQESGNNEVIFMALDLASLASVRAFATAFLSSEPRLDILIHNAG
ISSCGRTREAFNLLLRVNHIGPFLLTHTLLLPCLKACAPSRVVVVASAAHCRGRDLDFKRLDRP
VVGWRQELRAYADTKLANVLFARELANQLEATGVTCYAAHPGPVNSEFLRHVPGWLRPLLR
PLAWLVLRAPRGAQTPLYCALQEGIEPLSGRYFANCHVEEVPPAARDRAAHLWEASKRL
AGLGPGEDAEPDED PQSEDS EAPSSLS TPHP EETV SQPY SPQSSP DLSK MTHRIQAKVEP
EIQLS
```

Important features:

Signal peptide:

amino acids 1-16

Glycosaminoglycan attachment site.

amino acids 46-49

Short-chain alcohol dehydrogenase family

amino acids 37-49 and 114-124

FIGURE 76

GGAGGAGACAGCCTCTCGGGGGGAGGGGTTCCCTGCCTCTGCTGCTCCTGCTCATC**ATG**GGAGGATGGCTCAG
 GACTCCCCCCCCAGATCTAGTCCACCCCCAGGACAGCTGTTCCAGGGCCCTGGCCCTGCCAGGATGAGCTGC
 CAAGCTCAGGCCAGCCACCTCCACCATTCCGCTGGTTGCTGAATGGGCAGGCCCTGGAGCATGGTGCCTCCAGAC
 CCACACCACTCTCTGCTGATGGGACCCCTTCTGCTGCTACAGCCCCCTGCCCGGGACATGCCACGATGGCCAG
 GCCTGTCCACAGACTGGGTGTCTACACATGTAGGGCCAGCAACCGGCTTGGCAGCGAGTCCAGCAGAGGGCGCT
 CGCTCTCTGTGGCTGTCTCCCGGAGGATTTCCAGATCCAGCCTCGGGACATGGTGGCTGTGGTGGGTGAGCAG
 TTTACTCTGGAAATGGGGCGCCCTGGGGCCACCCAGAGCCACAGTCTCATGGTGGAAAGATGGGAAACCCCTG
 GCCTCCAGGCCGGAAGGCCACAGTGTCCGGGGGGTCCCTGCTGATGGCAAGACAGAGAAGATGACGAAGGG
 ACCTACATGTGTGGGCCACCAACAGCGCAGGACATAGGGAGAGCCGCGCAGCCGGGTTTCCATCCAGGAGCC
 CAGGACTACAGGGAGCTGTGGAGCTTCTGGCTGTGCGAACTCAGCTGGAAAGATGTGACATCTGCTGAACCCGGAT
 CCTGACAGAGGCCCAAGCCTAGACCGCGGGTGTGGCTCAGCTGGAAGTCCAGTGGCCCTGCTCGCTGCCCAA
 TCTTACACGGCCTTGTTCAGGACCCAGACTGCCCCGGAGGCCAGGGAGCTCCGTGGCAGAGGAGCTGCTGGCC
 GGCTGGCAGAGCGCAGAGCTTGGAGGCCCTCACTGGGGCCAAGACTACGAGTTCAAAGTGAGACCATCTCTGGC
 CGGGCTCAGAGCCCTGACAGCAACGTGCTGCTCTGAGGCTGCCGGAAAGATGCCAGTGGCCACCTCAGGAA
 GTGACTCTAAAGCTGGCAATGGCAGTGTCTTTGTGAGCTGGTCCCAACCACTGTGAAACCCCAATGGCATC
 ATCCGTGGCTACCAAGTCTGGAGCTTGGGCAACACATCACTGCCACAGCCCACTGGACTGTAGTTGGTGAGCAG
 ACCAGCTGGAAATCGCCACCCATATGCCAGGCTCTACTGTCGTGCAAGTGGCTCAGCTCACTGGTGTGGTGAGCAT
 GGGGAGCCCACTAGACTGTCTGTCTCTCTCTTTAGAGCAGGCCATGGAGCGAGGCCCAAGAACCTGAGTGGAGAT
 GGTCTCTGGACCTTGGAGCAGTGAAGGCTACTTGAAGCGGCTTGAGGTCACTGCCACCTGGCGTGTGGCACT
 TGGCTGCTGCTTCTGGGACCGCGGTGTGTATCCACCGCGCGCGCGAGCTTAGGTGGTCACTGGGCGGAGGTCTG
 TACAGATATACAGTGAAGATGCCATCTTAAACACAGGATGGATCAGAGTGACTCCCAAGTGGTGGGACAGCAT
 TGGCGTTCCACTCTGGCTCTCGGGACCTGAGCAGCAGCAGCCTCAGCAGTGGCTGGGGCGGATGCCCGG
 GAGCCACTAGACTGTGCTGCTCTTGTCTCTCTGGGACTCCCGAAGCCCGCGCTGGCCCTGCTTCCAGACACC
 AGCACTTTTTATGGCTCCCTCATGCTGAGTGGCTCCAGTACCCAGCCAGGCCAAGTCCCGAGTCCAGCT
 GTCAGGCGCTCCACCCCGCTGGGCCAGCTCTCCAGCCCTGTTCAGCTCAGACAGCTCTGCGAGCCGAGG
 GGACTCTCTCTCCCGCTTGTCTCTGGCCCTGCAGAGGCTTGAAGGCCAAAGAGCAGGAGCTGCAGCAT
 GCCAACAGTTCGCCACTGCTCCGGGGCAGCCACTCTTGGAGCTCCGGGCTGTGAGTTAGGAAATAGAGGTTCC
 AAGAACCTTTCCCAAAGCCAGGAGCTGTGCCCAAGCTCTGGTTGCTGGCGGGCTTGGGACCGAACTCTTC
 AGCTCTCTCAATGAGCTGGTTACTCGTCACTCTCCCTCCAGCACCCCTCTTCTCATGAACTCCCCAACTCAG
 AGTCAACAGACCCAGCTCCGGTGGCACCAAGGCTCCCTCCATCTCTGCTGAGCAGCAGCCCACTCCCATC
 CTTAGCCCTGCACTCCCCCTAGCCCCAGGCTCTCTCTCTGCCCCAGCCAGCTTCCAGTGCCTGTCT
 AGCTCTCTCACTGTCTCTCTGGGGAGGATCAAGACAGCTGTGACCCCTGAGGAGGTAGCCCTGTGCTTGGAA
 CTCAGTGAGGGTGAGGAGACTCCAGGAACAGCGTCTCTCCATGCCAAGGGCTCTTACCCCCCAACCACTAT
 GGGTACATCAGCGTCCCAACAGCCTCAGAGTTACGAGACATGGGCAGGACTGGAGGAGGGTGGGGCCAGGGG
 GGAGTCTTGTGTGCCACCTCGGCCCTGCTTACCCCCACCCAGCGAGGGCTCTTACCCATCTGTTGGGGC
 TCACTCTGAGGACAAATGCCAGCGCCAGAGCCAGCTTGTGAGTCTCCGATGGCTCTTCTGCTGCTGAT
 GCTCACTTTCGGCGGCCCTGGCAGTGGCTGTGGATAGCTTTGGTTCTGGTCTAGAGAGCCAGGGAGCAGATCG
 GTCTCTATAGATGCTCATCACTCTCCCTCCCAAGGGATGAGATCTTCTTCAAGCCCAACTCTCTCTCCCTGGCT
 TGGGATGGAGGCCAGACTGGTTGGAAGCATGGAGGTCAAGCCACACCCAGCGGCTGGGAAGGGGATGCCCTCC
 TGGCCCCCTGACTCTCAGATCTCTTCCAGAGAGACTCAGCTCCACTGATGTGGCCCAAGGCTGGTCTCTCTCT
 GTAGATTACTCT**GA**ACCGTGTCTCTGAGACTTCCAGACGGGAATCAGAACCATTCTCTCTGTCCACCCACAAG
 ACCTGGGCTGTGGTGTGGGCTTCTGGCGTGTGTTCTCTGCACTGGGGTCCACTTCCCAAGCTCTCCAGAGAG
 TTCTCTCTCCACGATTGTGAAACAAATGAAACAAATATAGACAAAGCTGACTGGAGCTCTCAGGGAGCAAA
 ACATCACTCCACCTGACTCTCTAGCCACTGCTTTCTCTCTGTGCCATCTCACTCCACACCAAGGTTGTTTGGC
 CTGAGGAGCAGCCCTGCCTGCTCTTCTCCCAACCATTGATCAGGAAGTGAGGAGCAGAGGTCCTTCT
 GTGGAGGACAGCAGTGGCTGTGGGAGAGGGCTGTGGAGGAAGGAGCTTCTCGGAGCCCACTCTCAGCTTACCT
 GGGCCCTCTCTTAGAGAAGAGCTCAACTCTCTCCAACTCACCATGGAAGAAATAATATGATGCCACTG
 AGGCTAGGAGCCCTACCTCATGCTCAAAACAAAGGCTTCAAGGCTGGGCTTAGCGAGGATGCTGAAGGAAGGAGG
 TATGAGACCTAGGCTCAAAGACCATCTCTCGTATGTTGTGCTATGAGCTTAGAGCTTAAGAAATTTGATACCAATAAT
 GGTAAAAA

FIGURE 77

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA41404
<subunit 1 of 1, 985 aa, 1 stop
<MW: 105336, pI: 6.55, NX(S/T): 7
MGGMAQDSPPPQILVHPQDQLFQGPGPARMSCQASGQPPPTIRWLLNGQPLSMVPPDPHLLP
DGTLLLLQPPARGHAHDQALSTDLGVYTCEASNRLGTAVSRGARLSVAVLREDFQIQPRDM
VAVVGEQFTLECGPPWGHPEPTVSWWKDGKPLALQGRHTVSGGSLLMARAEKSDEGTYMCV
ATNSAGHRESRAARVSIQEPQDYTEPVELLAVRIQLENVTLNLPDPAEGFKPRPAVWLSWKV
SGPAAPAQSYTALFRTQTAPGGQGAPWAEELLAGWQSAELGGLHWGQDYEFKVRPSSGRARG
PDSNVLLLRLEPEKVPSAPPQEVTLKPGNGTVFVSWVPPPAENHNGIIRGYQVWSLGNSTSLPP
ANWTVVGEQTQLEIATHMPGSYCVQVAAVTGAGAGEPSRPVCLLLEQAMERATQEPSEHGFW
TLEQLRATLKRPEVIATCGVALWLLLLGTAVCIHRRRRARVHLGPGLYRYSEDAILKHRMD
HSDSQWLADTWRSTSGSRDLSSSSSLSSRLGADARDPLDCRRSLLSWDSRSRPGVPLLPDTST
FYGSLIAELPSSTPARSPQVPAVRRLPPQLAQLSSPCSSSDSLCSRRGLSSPRLSLAPAEA
WKAKKQELQHANSPLLRGSHSLELRACELGNRGSKNLSQSPGAVPQALVAWRALGPKLLS
SSNELVTRHLPAPLFPHPETPPTQSQQTPPVAPQAPSSILLPAAPIPILSPCSPSPQASS
LSGSPSPASSRLSSSSSLSGEDQDSVLTPEEVALCLELSEGEETPRNSVSPMPRAPSPPTTY
GYISVPTASEFTDMGRTGGGVGPKGGVLLCPRPCLTPTPSEGLANGWGSASEDNAASARA
SLVSSSDGSFLADAHFARALAVAVDSFGFGLPREADCVFIDASSPPSPRDEIFLTPNLSLP
LWEWRPDWLEDMEVSHTQRLGRGMPPWPPDSQISSQRSQLHCRMPKAGASPVVDS

Important features:

Transmembrane domain:

amino acids 448-467

N-glycosylation sites:

amino acids 224-227, 338-341, 367-370, 374-377, 658-661 and 926-929

N-myristoylation sites.

amino acids 47-52, 80-85, 88-93, 99-104, 105-110, 181-186, 272-277, 290-295, 355-360, 403-408, 462-467, 561-566, 652-657, 849-854 and 876-881

Phosphotyrosine interaction domain proteins

amino acids 740-753

FIGURE 78

CTCCCACGGTGTCCAGCGCCCCAGAAATGCGGCTTCTGGTCTCTGCTATGGGGTTGCTCTGTGCT
CCCAGGTTATGAAGCCCTGGAGGGGCCAGAGGAAATCAGCGGTTTCGAAGGGGACACTGTGT
CCCTGCAGTGCACCTCAGGGGAAGAGCTGAGGGACCACCGGAAGTACTGGTGCAGGAAGGGT
GGGATCCTCTTCTCTCGTGTCTTGCCACCATCTATGCAGAAAGAAGGCCAGGAGACAAT
GAAGGGCAGGGGTGTCCATCCGTGACAGCCGCCAGGAGCTCTCGCTCATTGTGACCCCTGTGGA
ACCTCACCTGCAAGACGCTGGGGAGTACTGGTGTGGGTCGAAACCGGGGCCCGATGAG
TCTTTTACTGATCTCTGTTCGTCTTTCCAGGACCCTGCTGTCTCTCCCTCCCCTTCTCCCAC
CTTCCAGCCTCTGGCTACAACACGCTTGAGGCCAAAGGCAAAAGCTCAGCAAAACCCAGCCCC
CAGGATTGACTTCTCCTGGGCTCTACCCGCGAGCCACCAAGCCAGCAGGGGGAAGACAGGG
GCTGAGGCCCCCTCCATTGCCAGGACTTCCAGTACGGGCACGAAAGGACTTCTCAGTACAC
AGGAACCTCTCCTCACCCAGCGACCTCTCCTCTGACGGGAGCTCCCCCCCCCATGACAGC
TGGACTCCACCTCAGCAGAGGACACCAAGTCCAGCTCTCAGCAGTGGCAGCTCTAAGCCAGG
GTGTCCATCCCGATGGTCCGCATCTGCCCCAGTCTTGGTGTCTGCTGAGCCTTCTGTGAGC
CGCAGGCTGATCGCCTTCTGAGCCACCTGCTCTGTGGAGAAAGGAAGCTCAACAGGCCA
CGGAGACACAGAGGAACGAGAAGTTCTGGCTCTCAGCCTTGACTGCGGAGGAAAAGGAAGCC
CCTTCCAGGCCCTGAGGGGGACGTGATCTCGATGCTCTCCCTCCACACATCTGAGGAGGA
GCTGGGCTTCTCGAAGTTTGTCTCAGCGTAGGGCAGGAGGCCCTCTGGCCAGGCCAGT
GAAGCAGTATGGCTGGCTGGATCAGCACCGATTCCCGAAAGCTTTCACCTCAGCCTCAGAG
TCCAGTCTGCCCGACTCCAGGGCTCTCCACCTCTCCAGGCTCTCCTTGTGATGTTTCA
GCCTGACCTAGAAGCGTTTGTGAGCCCTGGAGCCAGAGCGGTGGCCTTGTCTTCCGGCTG
GAGACTGGGACATCCCTGATAGGTTACATCCCTGGGCAGAGTACCAGGCTGCTGACCCCTCA
GCAGGGCCAGACAAGGCTCAGTGGATCTGGTCTGAGTTTCAATCTGCCAGAACTCCTGGGC
CTCATGCCAGTGTGGAGCCCTGCCTTCTCCCACTCCAGACCCCACTTGTCTTCTCCTCCC
TGGCGTCTCAGACTTAGTCCACGCTCTCCTGCATCAGCTGGTGTGATGAAGAGGAGCATGT
GGGGTGAGACTGGGATTCTGGCTTCTCTTTGAACCCTGCATCCAGCCCTCAGGAAGCCT
GTGAAAAACGTGATTCTTGGCCCCACCAAGACCACCAAAACCATCTCTGGGCTTGGTGCAG
GACTCTGAATTCTAACATGCCAGTGAAGTGTGCACTTGTGAGTTTGGGGCCAGTGGGCCCTG
ATGAACGCTCAGCCCTTCTCAGCTTAGAGTCTGCATTGGGCTGTGACCTCTCCACCTGCC
CAATAGATCTGCTCTGTCTGCGACACAGATCCACGCTGGGGACTCCCTGAGGCGCTGCTAAG
TCCAGGCTTGGTCTGAGTCAAGTGCACATTGCGAGGATAAGCCAGGACCGGCACAGAAGTGG
TTGCCCTTNNCAATTTGCCCTCCCTGGNCCATGCCTTCTTGGCTTTGGAAAAAATGATGAAGA
AAACCTTGGCTCCTTCTTGTCTGGAAGGGTTACTTGCCTATGGGTTCTGGTGGCTAGAGA
GAAAAGTAGAAAAACAGAGTGCACGTAGGTGTCTAACACAGAGGAGAGTAGGAACAGGGCCG
ATACCTGAAGGTGACTCCGAGTCCAGCCCCCTGGAGAAGGGGTGGGGGTGGTGGTAAAGTA
GCACAACACTATTTTTTTCTTTTCCATTATTTGTTTTTAAAGACAGAACTCTCGTGTCT
GCTGCCAGGCTGGAGTGCAGTGGCACGATCTGCAAACTCCGCTCCTGGGTTCAAGTGATT
CTTCTGCTCAGCTCTCCGAGTAGCTGGGATTACAGGCACGCAACCAACACTGGCTGGTAAAT
TTTGTACTTTTGTAGATGGGGTTTCAACATGTTGGCCAGGCTGGTCTTGAACCTCTGAC
CTCAATGAGCCTCCTGCTTCACTCTCCAAATTCGCGGATTACAGGCATGAGCCACTGTG
TCTGGCCCTATTTCTTTAAAAAGTGAATTAAGAGTTGTTCAATGATGCAAACTTGGAAAG
ATGGAGGAGAAAAAGAAAAAGAAAAAATGTCAACATGCTCACCAGAGACTATCAT
TATTTCTGTTTGTGTAATCTCTTCACTCTTTCTTCTTCACTAATTTGCCGCTGTTCTT
TTTACAGAGCAATTAATCTGTATATACACTTTGTATCCTGCCTTTTCCACCTTATCGTTCC
ATCACTTTATTCAGCACTTCTGTGTTTTACAGACCTTTTATAAATAAATGTTTCATCA
GCTGCATAAAAAAAAAAAAAA

FIGURE 79

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44196
<subunit 1 of 1, 332 aa, 1 stop
<MW: 36143, pI: 5.89, NX(S/T): 1
MRLLVLLWGCLLLPGYEALEGPPEISGFEGDTVSLQCTYREELRDHRKYWCRKGGILFSRCS
GTIYAEEEGQETMKGRVSIIRDSRQELSLIVTLWNLTLDAGEYWCVGVEKRGPDSELLISLFV
FPGPCCPPSPSPPTFQPLATTRLQPKAKAQQTQPPGLTSPGLYPAATTAKQGKTGAEPPLPG
TSQYGHERTSQYTGTSPPHATSPAGSSRPPMQLDSTSAEDTSPALSSGSSKPRVSIIPMVRI
LAPVLVLLSLLSAAGLIAFCSHLLLWRKEAQQATETQRNEKFWLSRLTAEKEAPSQAPEGD
VISMPPLHTSEELGFSKFVSA
```

Important features:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 248-269

N-glycosylation site.

amino acids 96-99

Fibrinogen beta and gamma chains C-terminal domain.

amino acids 104-113

Ig like V-type domain:

amino acids 13-128

FIGURE 80

TTGTGACTAAAAGCTGGCCTAGCAGGCCAGGGAGTGCAGCTGCAGGCGTGGGGGTGGCAGGA
GCCGCAGAGCCAGAGCAGACAGCCGAGAAAACAGGTGGACAGTGTGAAAGAACCAGTGGTCTC
GCTCTGTTGCCCAGGCTAGAGTGTACTGGCGTGATCATAGCTCACTGCAGCCTCAGACTCCT
GGACTTGAGAAATCCTCCTGCCTTAGCCTCCTGCATATCTGGGACTCCAGGGGTGCACTCAA
GCCCTGTTTCTTCTCCTTCTGTGAGTGGACCACGGAGGCTGGTGAGCTGCCTGTTCATCCCAA
AGCTCAGCTCTGAGCCAGAGTGGTGGTGGCTCCACCTCTGCCGCCGGCATAGAAGCCAGGAG
CAGGGCTCTCAGAAGGCGGTGGTGGCCAGCTGGGATCATGTTGTTGGCCCTGGTCTGTCTGC
TCAGCTGCCTGCTACCTCCAGTGAGGCCAAGCTCTACGGTCGTTGTGAACTGGCCAGAGTG
CTACATGACTTCGGGCTGGACGGATACCGGGGATACAGCCTGGCTGACTGGGTCTGCCTTGC
TTATTTCAACAAGCGGTTTCAACGCAGCTGCTTTGGACTACGAGGCTGATGGGAGCACCAACA
ACGGGATCTTCAGATCAACAGCCGGAGGTGGTGCAGCAACCTCACCCCGAACGTCCCCAAC
GTGTGCCGGATGTACTGCTCAGATTTGTTGAATCCTAATCTCAAGGATACCGTTATCTGTGC
CATGAAGATAACCCAGAGCCTCAGGGTCTGGGTTACTGGGAGGCCCTGGAGGCATCACTGCC
AGGGAAAAGACCTCACTGAATGGTGGATGGCTGTGACTTCTTAGGATGGACGGAACCATGCA
CAGCAGGCTGGGAAATGTGGTTTGGTTCCTGACCTAGGCTTGGGAAGACAAGCCAGCGAATA
AAGGATGGTTGAACGTGAAA

FIGURE 81

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52187

<subunit 1 of 1, 146 aa, 1 stop

<MW: 16430, pI: 5.05, NX(S/T): 1

MLLALVCLLSCLLPSSSEAKLYGRCELARVLHDFGLDGYRGYSLADWVCLAYFTSGFNAAALD
YEADGSTNNGIFQINSRRWCSNLTPNVPNVCRMVCSDLLNPNLKDTVICAMKITQEPQGLGY
WEAWRHHCQKDLTEWVDGPDF

Important features:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 67-72

Homologous region to Alpha-lactalbumin / lysozyme C proteins.

amino acids 34-58 (catalytic domain), 111-132 and 66-107

FIGURE 82

AGCCGCTGCCCGGGCCGGGCGCCCGCGCGGCACCATGAGTCCCCGCTCGTGCTCGGTTT
GCTGCGCCTCCTCGTCTTCGCGCTCTTCTCAGCCGCGCGAGCAACTGGCTGTACCTGGCCCA
AGCTGTCGTCGGTGGGGAGCATCTCAGAGGAGGAGACGTGCGAGAACTCAAGGGCCTGATC
CAGAGGCAGGTGCAGATGTGCAAGCGGAACCTGGAAGTCATGGACTCGGTGCGCCGCGGTGC
CCAGCTGGCCATTGAGGAGTGCCAGTACCAGTTCGGAACCGCGCTGGAACTGCTCCACAC
TCGACTCCTTGCCCGTCTTCGGCAAGGTGGTGACGCAAGGGAAGTTCGGAGGCGGCCTTCGTG
TACGCCATCTCTTCGGCAGGTGTGGCCTTTGAGTACGCGGGCGTGTCAGCAGTGGGGAGCT
GGAGAAGTGCGGTGTGACAGGACAGTGCATGGGGTCAGCCACAGGGCTTCAGTGGTCAG
GATGCTCTGACAACATCGCCTACGGTGTGGCCTTCTCACAGTCGTTTGTGGATGTGCGGGAG
AGAAGCAAGGGGGCCTCGTCCAGCAGAGCCCTCATGAACCTCCACAACAATGAGGCCGGCAG
GAAGGCCATCCTGACACACATGCGGGTGAATGCAAGTGCCACGGGGTGTGAGGCTCCTGTG
AGGTAAGACGTGCTGGCGAGCCGTGCGCCCTTCCGCCAGGTGGGTACGCACTGAAGGAG
AAGTTTGTAGTGCCACTGAGGTGGAGCCACGCCGCTGGGCTCCTCCAGGGCACTGGTACC
ACGCAACGCACAGTTCAAGCCGCACACAGATGAGGACCTGGTGTACTTGGAGCCTAGCCCCG
ACTTCTGTGAGCAGACATGCGCAGCGCGTGTGGGCAGGAGGGCCGCACATGCAACAAG
ACGTCCAAGGCCATCGACGGCTGTGAGCTGCTGTGCTGTGGCCGCGGCTTCCACAGGCGCA
GGTGGAGCTGGCTGAACGCTGACGTGCAAATTCACATGGTGTGCTTGTCAAGTGCCGGC
AGTGCCAGCGGCTCGTGGAGTTGCACAGTGCAGATGACCGCTGCTAGCCCTGCGCCCGGC
AACCACCTAGTGGCCAGGGAAGGCCGATAATTTAAACAGTCTCCACCACCTACCCCAAGA
GATACTGGTTGTATTTTGTCTGTTTGGTTTGGGTCTCATGTTATTTATTGCCGAA
ACCAGGCAGGCAACCCCAAGGGCACCAACCAGGGCCTCCCCAAAGCCTGGGCCTTTGTGGCT
GCCACTGACCAAAGGGACCTTGCTCGTGGCGCTGGCTGCCCGCATGTGGCTGCCACTGACCA
CTCAGTTGTTATCTGTGTCGCTTTTCTACTTGACAGACCTAAGGTGGAGTAACAAGGAGTAT
TACCACCACATGGCTACTGACCGTGTATCGGGGAAGAGGGGGCCTTATGGCAGGGAAAATA
GGTACCGACTTGATGGAAGTCACACCCTCTGGAAAAAGAACTTTAACTCTCCAGCACACA
TACACATGGACTCCTGGCAGCTTGAGCCTAGAAGCCATGTCTCTCAAATGCCCTGAGAAAGG
GAACAAGCAGATACAGGTCAAGGGCACAGGTTCAATTTAGCCCTTACATGGACAGCTAGA
GGTTGCATATCTGTGGGTCTTCCAGGCAAGAAGGGGAGATGAGAGCAAGAGACGACTGAA
GTCCACCCCTAGAACCAGCCTGCCCCAGCCTGCCCCCTGGGAAGAGGAAACTTAACCACTCC
CCAGACCCACCTAGGCAGGCATATAGGCTGCCATCCTGGACCAGGGATCCCGGCTGTGCCTT
TGCAGTCATGCCCCAGTCACTTTTACAGCGCTGTTCTCCATGAAACTGAAAAACACACAC
ACCTGCGAGA
GAGAGGGAGGAAAGGGCTGTGCCTTTGAGTCATGCCGAGTCACCTTTCACAGCACTGTTCTCT

FIGURE 83

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48328
<subunit 1 of 1, 351 aa, 1 stop
<MW: 39052, pI: 8.97, NX(S/T): 2
MSPRSCLSRLRLLVFAVFSAAASNWLYLAKLSSVGSISEEETCEKLKGLIQRQVMCKRNLE
VMDSVRRGAQLAIEECQYQFRNRWNCSTLDSLVPVFGKVVTQGTREAAFVYAISSAGVAFV
TRACSSGELEKCGCDRTVHGVSPQGFQWSGCSDNIAYGVAFSQSFVDVRERSKGASSRALM
NLHNNEAGRKAILTHMRVECKCHGVSGSCEVKTCWRAVPPFRQVGHALKEKFDGATEVEPRR
VGSSRALVPRNAQFKPHTDEDLVYLEPSPDFCEQDMRSGVLGTRGRTCNKTSKAIDGCELLC
CGRGFHTAQVELAERCSCKFHWCCFVKCRQCQRLVELHTCR
```

Important features:

Signal peptide:

amino acids 1-22

N-glycosylation sites.

amino acids 88-91 and 297-300

Wnt-1 family signature.

amino acids 206-215

Homologous region to Wnt-1 family proteins

amino acids 183-235, 305-350, 97-138, 53-92 and 150 -174

FIGURE 84

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCTGGGTGCCTGCAT
CGCCATGGGACACCACCAGGTACAGCAAGTGGGCGGCAGCTCCGAGGAGTCCCCGGAGGGC
CCTGGGGACGCTGGGTGCACTGGAGCAGGAGACCCCTCTTCTTGCCCTGGCTGTCTTGCT
ACCACAGTCTTTGGGCTGTGATTCTGAGTATCCTATTGTCCAAGGCCTCCACGGAGCGCGC
GGCGCTGCTTGACGGCCACGACCTGCTGAGGACAAACGCCTCGAAGCAGACGGCGGCGCTGG
GTGCCCTGAAGGAGGAGGTGCGAGACTGCCACAGCTGCTGCTCGGGGACGCAGGCGCAGCTG
CAGACCACGCGCGCGGAGCTTGGGGAGGCGCAGGCGAAGCTGATGGAGCAGGAGAGCGCCCT
GCGGGAACTGCGTGAGCGCGTGACCCAGGGCTTGGCTGAAGCCGGCAGGGGCCGTGAGGACG
TCCGCACTGAGCTGTTCCGGGCGCTGAGGCCGTGAGGCTCCAGAACAACCTCTGCGAGCCG
TGCCCCACGTGCTGGCTGTCTTTCGAGGGCTCCTGCTACTTTTTCTCTGTGCCAAAGACGAC
GTGGGCGGCGGCGCAGGATCACTGCGCAGATGCCAGCGCGCACCTGGTGATCGTTGGGGGCC
TGGATGAGCAGGGCTTCTCACTCGGAACACGCGTGGCCGTGGTTACTGGCTGGGCTGAGG
GCTGTGCGCCATCTGGGCAAGGTTCAAGGCTACCACTGGGTGGACGGAGTCTCTCTCAGCTT
CAGCCACTGGAACCAGGAGAGCCCAATGACGCTTGGGGGCGCGAGAAGTGTGTCTATGAGC
TGACACGCGGGCTGTGGAACGACGACCCGTGTGACAGCGAGAAGGACGGCTGGATCTGTGAG
AAAAGGCACAACCTGCTTGACCCCGCCAGTGCCCTGGAGCCGCGCCATTGCAGCATGTCGTA
TCCTGGGGCTGCTCACCTCCCTGGCTCCTGGAGCTGATTGCCAAGAGTTTTTTCTTCTCCT
CATCCACCCTGCTGAGTCTCAGAAACACTTGGCCCAACATAGCCCTGTCCAGCCCAAGTGCC
TGGGCTCTGGGACCTCCATGCCGACCTCATCCTAACTCCACTCAGCGAGACCCAACTAACC
TCCACTAGCTCCAAAATCCCTGCTCCTGCGTCCCCGTGATATGCCTCCACTTCTCTCCCTAA
CCAAGGTTAGGTGACTGAGGACTGGAGCTGTTTGGTTTTCTCGCATTTTCCACCAAAGTGA
AGCTGTTTTTGACGCTGAGGAAGCATCAATAAATATTTGAGAAATGAAAAA

FIGURE 85

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56352
<subunit 1 of 1, 293 aa, 1 stop
<MW: 32562, pI: 6.53, NX(S/T): 2
MDTTRYSKWGSSEEVPGGPWGRVHVHWSRRPLFLALAVLVTTVLWAVILSILLSKASTERAA
LLDGHDLLRTNASKQTAALGALKEEVGDCHSCSGTQAQLQTTRAELGEAQAKLMEQESALR
ELRERVVTQGLAEAGRGREVDVTELFRALAVRLQNNSCPCPTSWLSFEGSCYFFSVPKTTW
AAAQDHCADASAHLVIVGGLDEQGFLTRNTRGRGYWLGLRAVRHLGKVQGYQWVDGVSLSPS
HWNQGEPNDAWGRENVMMLHTGLWNDAPCDSEKDGWICEKRHNC

Important features:

Type II transmembrane domain:

amino acids 31-54

N-glycosylation sites.

amino acids 73-76 and 159-162

Leucine zipper pattern.

amino acids 102-123

N-myristoylation sites.

amino acids 18-23, 133-138 and 242-247

C-type lectin domain signature.

amino acids 264-287

FIGURE 86

GCCAGGGGAAGAGGGTGATCCGACCCGGGGAAGGTCGCTGGGCAGGGCGAGTTGGGAAAGCG
GCAGCCCCCGCCGCCCCCGCAGCCCCCTTCTCCTCCTTTCTCCACGTCTATCTGCCTCTCG
CTGGAGGCCAGGCCGTGCAGCATCGAAGACAGGAGGAACCTGGAGCCTCATTGGCCGGCCCCGG
GGCGCCGGCCTCGGGCTTAAATAGGAGCTCCGGGCTCTGGCTGGGACCCGACCGCTGCCGGC
CGCGCTCCCGCTGCTCCTGCCGGGTGATGGAAAAACCCAGCCCCGGCCGCCCTGGGCAAG
GCCCTCTGCGCTCTCCTCTGGCCACTCTCGGCGCCGCCGGCCAGCCTCTTGGGGGAGAGTCT
CATCTGTTCCGCCAGAGCCCCGGCCAAATACAGCATCACCTTCACGGGCAAGTGGAGCCAGA
CGGCCTTCCCCAAAGCAGTACCCCCCTGTTCCGCCCCCTTGGCAGTGGTCTTCGCTGCTGGGG
GCCGCGCATAGCTCCGACTACAGCATGTGGAGGAAGAACCAGTACGTAGTAAACGGGCTGCG
CGACTTTGCGGAGCGCGGCGAGGCCCTGGGCGCTGATGAAGGAGATCGAGGCGCGGGGGAGG
CGCTGCAGAGCGTGACGAGGTGTTTTCGGCGCCCGCGCTCCCGAGCGCACCGGGCAGAGC
TCGGCGGAGCTGGAGGTGCAGCGCAGGCACCTCGCTGGTCTCGCTTGTGGTGCGCATCGTGCC
CAGCCCCGACTGGTTCTGTGGGCGTGGACAGCCTGGACCTGTGCGACGGGGACCGTTGGCGGG
AACAGGCGGGCGTGGACCTGTACCCCTACGACGCGGGGACGGACAGCGGCTTACCTTCTCC
TCCCCAACTTCGCCACCATCCCGCAGGACACGGTGACCGAGATAACGTCTCCTCTCCAG
CCACCCGGCCAACTCCTTCTACTACCCGCGGCTGAAGGCCCTGCCTCCCATCGCCAGGGTGA
CACTGCTGCGGCTGCGACAGAGCCCCAGGGCCTTCATCCCTCCCGCCCCAGTCTGCCCCAGC
AGGGACAATGAGATTGTAGACAGCGCCTCAGTTCAGAAACGCCGCTGGAATGCGAGGTCTC
CCTGTGGTCGTCTTGGGACTGTGCGGAGGCCACTGTGGGAGGCTCGGGACCAAGAGCAGGA
CTCGCTACGTCCGGGTCCAGCCCGCCAACAACGGGAGCCCTGCCCGAGCTCGAAGAAGAG
GCTGAGTGGCTCCCTGATAACTGCGCTCTAAGACCAGAGCCCCGAGCCCCCTGGGGCCCCCG
GAGCCATGGGGTGTGGGGGCTCCTGTGCAGGCTCATGCTGCAGGCGGCCGAGGGCACAGGG
GGTTTCGCGCTGCTCCTGACCGCGGTGAGGCCGCGCCGACCATCTCTGCACTGAAGGGCCCT
CTGGTGGCCGGCACGGGCATTGGGAAACAGCCTCCTCCTTCCCAACCTTGCTTCTTAGGGG
CCCCCGTGTCCCGTCTGCTCTCAGCCTCCTCCTGTCAGGATAAAGTCATCCCCAAGGCTC
CAGCTACTCTAAATTATGTCTCCTTATAAGTTATTGCTGTCCAGGAGATTGTCTTTCATCG
TCCAGGGGCTGGCTCCCACGTGGTTGCAGATACCTCAGACCTGGTGTCTAGGCTGTGCTG
AGCCCACTCTCCCGAGGGCGCATCAAGCGGGGGCCACTTGAGAAGTGAATAAATGGGGCGG
TTTCGGAAGCGTCAGTGTTTCCATGTTATGGATCTCTCTGCGTTTGAATAAAGACTATCTCT
GTTGCTCACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 87

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53971

><subunit 1 of 1, 331 aa, 1 stop

><MW: 35844, pI: 5.45, NX(S/T): 2

MENPSPAAALGKALCALLLATLGAAGQPLGGESICSARAPAKYSITFTGKWSQTAFPKQYPL
FRPPAQWSSLLGAAHSSDYSMWRKNQYVSNGLRDFAEERGEAWALMKEIEAAGEALQSVHEVF
SAPAVPSGTGQTSAELEVQRHSLVSFVVRIVPSPDWFGVDSLDCDGDWRREQAALDLYP
YDAGTDSGFTFSSPNFATIPQDTVTEITSSSPSHPANFYYPRLKALPPIARVTLLRLRQSP
RAFIPPAPVLPSRDNEIVDSASVPETPLDCEVSLWSSWGLCGGHCGRLGTSRTRYVRVQPA
NNGSPCPELEEEAECPDNCV

Important features:

Signal peptide:

amino acids 1-26

FIGURE 88

GGCGGCGTCCGTGAGGGGCTCCTTTGGGCAGGGGTAGTGTGTTGGTGTCCCTGTCTTGCCTGA
TATTGACAAACTGAAGCTTTCTGCACTGACTTAAGGAAGAGTGTACTCGTAGGCGGA
CAGCTTTAGTGGCCGGCCGGCCGCTCTCATCCCCGTAAGGAGCAGAGTCCTTTGTACTGAC
CAAGATGAGCAACATCTACATCCAGGAGCCTCCACGAATGGGAAGGTTTTATTGAAAACTA
CAGCTGGAGATATTGACATAGAGTTGTGGTCCAAAGAAGCTCCTAAAGCTTGCAAGATTTT
ATCCAACCTTTGTTTGAAGCTTATTATGACAATACCATTTTTCATAGAGTTGTGCCTGGTTT
CATAGTCCAAGGCGGAGATCTACTGGCACAGGAGTGGTGGAGAGTCTATCTATGGAGCGC
CATTCAAAGATGAATTCATTACGGTTGCGTTTTAATCGGAGAGGACTGGTTGCCATGGCA
AATGTGGTTCTCATGATAATGGCAGCCAGTTTTCTTCACACTGGGTGAGCAGATGAAC
TAACAATAAGCATACCATCTTTGGAAGGTTACAGGGGATACAGTATATAACATGTTGCGAC
TGTGAGAAGTAGACATTGATGATGACGAAAGACCACATAATCCACACAAAAATAAAAGCTGT
GAGGTTTTGTTTAACTCCTTTTGATGACATCATTCGAAGGAAATTAAGGCTGAAAAAAGA
GAAACCGAGGAGGAAGTAAAGAAATGAAACCCAAGGCACAAAAAATTTAGTTTACTTT
CATTTGGAGAGGAAGCTGAGGAAGAAGAGGAGGAAGTAAATCGAGTTAGTCAGAGCATGAAG
GGCAAAAGCAAAAGTAGTCATGACTTGCTTAAGGATGATCCACATCTCAGTTCTGTTCCAGT
TGTAAGAGTGAAGAGGTGATGCACAGATTTAGTTGATGATGGAGAAGATGAAAGTGCAG
AGCATGATGAATATATTGATGGTGATGAAAGAACCCTGATGAGAGAAAGAAATTGCCAAAAA
TTAAAAAAGGACACAAGTGCGAATGTTAAATCAGCTGGAGAAGGAGAAGTGGAGAAGAAATC
AGTCAGCCGAGTGAAGAGCTCAGAAAAGAAGCAAGACAATTAACCGGAACCTTAGCAG
CAAAACAAAAAAGTAGAAATGCAGCAAAACAAGCAGAAAAAAGAAGTGAAGAGGAAGAA
GCCCTCCAGATGGTGTGTTGCCGAATACAGAAGAGAAAAGCAAAAGTATGAAGCTTTGAG
GAAGCAACAGTCAAAGAAGGGAACCTCCCGGAAGATCAGACCCTTGCACTGCTGAACCAGT
TTAAATCTAACTCACTCAAGCAATTGCTGAAACACCTGAAATGACATTCCTGAAACAGAA
GTAGAAGATGATGAAGGATGGATGTCACATGTACTTCAGTTTGAGGATAAAGCAGAAAAGT
GAAAGATGCAAGCATGCAAGACTCAGATACATTGAAATCTATGATCCTCGGAATCCAGTGA
ATAAAAGAAGGAGGGAAGAAAGCAAAAGCTGATGAGAGAGAAAAAGAAAGAAGATAAAT
GAGAATAATGATAACAGAACCTTGCTGGAATGTGCCTACAATGGCCTTGAACAGCCATTG
TTCCCAACAGCATCACTTAGGGGTGTGAAAGAAGTATTTTTGAACCTGTTGCTGGTTTTG
AAAAACAATTATCTGTTTTGCAAAATGTGGAATGATGTAAGCAATGCTTTTGGTTACTGG
TACATGTGTTTTTTCCTAGCTGACCTTTTATATTGCTAAATCTGAAATAAAATAAATTTCTT
TCCACAAAAA

FIGURE 89

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50919
><subunit 1 of 1, 472 aa, 1 stop
><MW: 53847, pI: 5.75, NX(S/T): 2
MSNIYIQEPPTNGKVLLKTTAGDIDIELWSKEAPKACRNFIQLCLEAYDNTIFHRVVPGFI
VQGGDPTGTGSGGESIYGAPFKDEFHSRLRFNRRGLVAMANAGSHDNGSQFFFTLGRADELN
NKHTIFGKVTGDTVYNMLRLSEVDIDDDERPHNPHKIKSCEVLFPNPFDDIIPREIKRLKKEK
PEEEVKKLKPKGTTKNFSLLSFGEEAESEEEVNRVSQSMKGKSKSSHDLKDDPHLSSVPVV
ESEKGDAPDLVDDGEDESAEHDEYIDGDEKNLMRERIAKLLKDDTSANVKSAGEGEVEKKSV
SRSEELRKEARQLKRELLAAKQKKVENAAKQAEKRSEEEAEPPDGAVAEYRREKQKYEALRK
QQSKKGTSTREDQTLALLNQFKSKLTQAIATPENDIPETEVEDDEGWMShVLQFEDKSRKVK
DASMQDSDTFEIIYDPRNPVNKRREESKKLMREKKERR
```

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 109-112 and 201-204

Cyclophilin-type peptidyl-prolyl cis-trans isomerase signature.

amino acids 49-66

Homologous region to Cyclophilin-type peptidyl-prolyl cis-trans isomerase

amino acids 96-140, 49-89 and 22-51

FIGURE 90

CGCCGCCGTTGGGGCTGGAAGTTCCCGCCAGGTCCGTGCCGGGCGAGAGAGATGCTGCCCGG
 CCCGCTCGGCCTTTGAGGCGAGAGAAGTGTCCAGACCCATTTCCGCTTGCTGACGGCGTCC
 AGCCCTGGCCAGACATGTCCACAGGGTTCTCTTCGGGTCCGGGACTCTGGGCTCCACCAC
 GTGGCCGCCGGGGGACCAGCACAGGCGCGTCTTTCTCCTTCGGGAACGGGAACGTCTAGCAA
 CCCCTCTGTGGGGCTCAATTTTGGAAACTTTGGAAGTACTTCAACTCCAGCAACTACATCTG
 CTCCTTCAAGTGGTTTTGGAAACCGGGCTCTTTGGATCTAAACCTGCCACTGGGTTCACTCTA
 GGAGGAACAAATACAGGTGCCTTGACACCAAGAGGCTCAAGTGGTACCAAATATGGAAC
 CCTGCAAGGAAAAACAGATGTCATGTGGGGAAGACACCCATCCAAGTCTTTTATAGGAGTCCCTT
 TCTCCAGACCTCTCTTAGGTATCTCAGGTTTGCACCTCCAGAACCCCGGAGCCCTGGAAA
 GGAATCAGAGATGCTACCACCTACCCGCTGGATGGAGTCTCGCTCTGTGCGCCAGGCTGGAG
 TGCAGTGGCAGCATCTCGGCTCACTGCAACCTCCGCTCCCGGGTCAAGCGAGTCTCTGCG
 CTGAGCCTCTGAGTGTCTGGGGCTACAGGTGCTGACAGGAGTCTGGGGCCAGCTGGCCCTG
 ATGTACGTGAGCACGCGGGAAACGGTACAAGTGGCTGCGCTTCAGCGAGGACTGTCTGTACCT
 GAACGTGTACGCGCCGCGCGCGCCCGGGATCCCGAGCTGCCAGTATGGTCTGCTGTTCC
 CGGGAGGCGCCTTCATCGTGGGCGCTGCTTCTTCGTACGAGGGCTCTGACTTGGCCGCCCGC
 GAGAAAGTGGTCTGTTGTTCTGACGACACAGGCTCGGCATCTTCGGCTTCTGAGCACGGA
 CGACAGCCAACGCGCGCGGGAACCTGGGGGCTGCTGGACAGATGGCGGCTCTGCGCTGGGTGC
 AGGAGAACATCGCAGCCTTCGGGGGAGACCCAGGAATGTGACCTGTTCGGCCAGTCCGCG
 GGGGCCATGAGCATCTCAGGACTGATGATGTACCCCTAGCCTCGGGTCTCTTCCATCGGGC
 CATTTCACAGAGTGGCACCGGTTATTACAGACTTTTATCATCTAGTAACCCACTGAAAGTGG
 CCAAGAAGGTGCGCCACTGGCTGGATGCAACCAACAGCACACAGATCCTGGTAAACTGC
 CTGAGGGCACTATCAGGGACCAAGGTGATGCGTGTGTCCAACAAGATGAGATTCTCCAACT
 GAACCTCCAGAGAGACCCGGAAGAGATTATCTGGTCCATGAGCCTGTGGTGGATGGTGTGG
 TGATCCCAGATGACCTTTGGTGTCTGACCCAGGGGAAGGTTTTCATCTGTGCCCTACCTT
 CTAGGTGTCAACAACCTGGAATTCAATTGGCTCTTGCTTTATAATATCAACAAGGAGCAGT
 ACCACTTGTGGTGGAGGAGTACCTGGACAATGTCAATGAGCATGACTGGAAGATGTACGAA
 ACCGTATGATGGACATAGTTCAAGATGCCACTTTCTGTATGCCACACTGCAGACTGCTCAC
 TACCACCGAATAACCCCAATGATGGGAATCTGCCCTGTGGCCACGCTACAACAAGGATGAA
 AAGTACCTGCAGCTGGATTTTACCACAAGAGTGGGCTAGAAAGCTCAAGGAGAAGAAGATGGC
 TTTTGGATGAGTCTGTACCACTCTCAAAGACCTGAGAAGCAGAGGCCAATTCTAAGGGTGGC
 TATGCAGGAAGGAGCCAAAGAGGGGTTTCCCCCACCATCCAGGCCCTGGGGAGACTAGCCA
 TGGACATACCTGGGGACAAGATTTACCCACCCAGTTTAGAATGCAAGGAGCTCCCTGCT
 GCCTCAGGCGCAAAGCTAGAGCTTTTGCTGTGTGTGGGACCTGCAGTGCCTTTCCAGCC
 TGACATCCCATGATGCCCTCTACTTCACTGTTGACATCCAGTTAGGCCAGGCCCTGTCAAC
 ACCACACTGTGCTCAGCTCTCCAGCCTCAGGACAACCTCTTTTTTCCCTTCTTCAAATCTCT
 CCCACCTTCAATGTCTCTTGTGACTCCTTCTTATGGGAGGTGCAGCCAGACTGCCACTGC
 CCCTGTCACTGCACCCAGCTTGGCATTACCATCCATCTGCTCAACCTTGTTCCTGTCTGT
 TCACATTGGCCTGGAGGCTTAGGGCAGGTTGTGACATGGAGCAACTTTGGTAGTTTGGGA
 TCTTCTCTCCACCCACACTTATCTCCCCAGGGCCACTCCAAGTCTATACACAGGGGTGG
 TCTCTTCAATAAAGAAGTGTGATTAGAAAAA

FIGURE 91

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44179
<subunit 1 of 1, 545 aa, 1 stop
<MW: 58934, pI: 9.45, NX(S/T): 4
MSTGFSFGSGTLGSTTVAAGGTSTGGVFSFGTGTSSNPSVGLNFGNLGSTSTPATTAPSSG
FGTGLFGSKPATGFTLLGGTNTGALHTKRQVVTKYGTLQGKQMHVGKTFIQVFLGVFFSRPP
LGILRFAPPEPPEPWKGIRDATTYPGW$LALSPGWSAVARSRLTATSASRVQASLLPQPLS
VWGYRCLQESWGQLASMYVSTRERYKWLRFSEDCLYLNVIYAPARAPGDPQLPVMVWFPGGAF
IVGAASSYEGSDLAAREKVVLVFLQHRLGIFGFLSTDDSHARGNWGLLDQMAALRWVQENIA
AFGGDPGNVTLFGQSAGAMSISGLMMSPLASGLFHRAISQSGTALFRLFITSNPLKVAKKVA
HLAGCNHNSTQILVNCRLALSGTKVMRVSNKMRFLQLNFQRDP EEI IWSMSPVVDGVVIPDD
PLVLLTQGKVSSVPYLLGVNNLEFNWLLFPYNITKEQVPLVVEEYLDNVNEHDWKMLRNRMD
IVQDATFVYATLQTAHYHRETMMGICPAGHATTRMKSTCSWILPQEWA
```

Important features:

Signal peptide:

amino acids 1-29

Carboxylesterases type-B serine active site.

amino acids 312-327

Carboxylesterases type-B signature 2.

amino acids 218-228

N-glycosylation sites.

amino acids 318-321, 380-383 and 465-468

FIGURE 92

GAGAACAGGCCTGTCTCAGGCAGGCCCTGCGCCTCCTATGCGGAGATGCTACTGCCACTGCT
GCTGTCCTCGCTGCTGGGCGGGTCCCAGGCTATGGATGGGAGATTCTGGATACGAGTGCAGG
AGTCAGTGATGGTGCCGGAGGGCCCTGTGCATCTCTGTGCCCTGCTCTTTCTCCTACCCCCGA
CAAGACTGGACAGGGTCTACCCAGCTTATGGCTACTGGTTCAAAGCAGTGACTGAGACAAC
CAAGGGTGTCTCTGTGGCCACAAACCACAGAGTTCGAGAGGTGGAAATGAGCACCCGGGGCC
GATTCAGCTCACTGGGGATCCCGCAAGGGGAAGTCTCTTGGTGATCAGAGCACGCGCAG
ATGCAGGATGAGTCACAGTACTTCTTTCCGGTGGAGAGGGAAGCTATGTGACATATAATTT
CATGAACGATGGGTCTCTTTCTAAAGTAACAGTGTCTCAGCTTCAGCCCCAGACCCAGGACC
ACAACCCAGACCTCACTGCCATGTGGACTTCTCCAGAAAGGGTGTGAGCGCACAGAGGACC
GTCCGACTCCGTGTGGCTATGCCCCAGAGACCTTGTATCAGCATTTCACTGTGACAACAC
GCCAGCCCTGGAGCCCCAGCCCCAGGGAAATGTCCCATACCTGGAAGCCCCAAAAGGCCAGT
TCCTGCGGCTCCTCTGTGTGCTGACAGCCAGCCCCCTGCCACACTGAGCTGGGTCTCTGCAG
AACAGAGTCTCTCCTCGTCCCATCCCTGGGGCCCTAGACCCCTGGGGCTGGAGCTGCCCGG
GGTGAAAGCTGGGATTCAGGCGCTACACCTGCCGAGCGGAGAACAGGCTTGGCTCCCAGC
AGCGAGCCCTGGACCTCTCTGTGCAGTATCTCTCCAGAGAACCCTGAGAGTGATGGTTTCCCAA
GCAACACGAGCAGTCTCTGGAAAACCTTGGGAACGGCAGCTCTCTCCAGTACTGGAGGGGCCA
AAGCTGTGCCCTGGTCTGTGTACACACAGCAGCCCCCAGCAGGCTGAGCTGGACCCAGA
GGGGACAGTTCTGAGGCCCTCCCAGCCCTCAGACCCCGGGGTCTCTGGAGCTGCCCTCGGGTT
CAAGTGGAGCACGAAGGAGAGTTTACCTGCCACGCTCGGCACCCACTGGGCTCCCAGCAGCT
CTCTCTCAGCCTCTCCGTGCACATAAGAAAGGAGTCACTCTCAACGGCTACTCTCCAGCGAG
CGTTTTCTGGGAATCGGCATCACGGCTCTTCTTTCTCTGCTGGCCCTGATCATCATGAAG
ATTCTACCGAAGACGGACTCAGACAGAAAACCCGAGGCCAGGTTCTCCCGGCACAGCAC
GATCTCGGATTACATCAATGTGCTCCCGACGGCTGGCCCCCTGGCTCAGAAGCGGAATCAGA
AAGCCACACCAACAGTCTCGGACCCCTCCTCCACAGGCTGCTCCTCCCGAGAATCAAAG
AAGAACCGAAGAAAGCAGTATCAGTTGCCAGTTTCCAGAACCCAAATCATCCACTCAAGC
CCCAGAATCCCAGGAGAGCCAAGAGGAGCTCCATTATGCCACGCTCAACTTCCAGGCGCTCA
GACCCAGCCCTGAGGCCCGGATGCCCAAGGGCACCCAGGCGGATTATGCAGAAGTCAAGTTT
CAATGAGGGTCTCTTAGGCTTTAGGACTGGGACTTCGGCTAGGGAGGAAGGTAGAGTAAGAG
GTTGAAGATAACAGAGTGCAAGTTCCTTCTCTCCTCTCTCTCTCTCTCTCTCTCTCTCTCT
CTCTCTTTCTCTCTTTTAAAAAAACATCTGGCCAGGGGCACATGCTGCTCAGCGCTGTAATC
CCAGCACTTGGGAGGTTGAGGTTGGGCAGATCGCCTGAGGTCTGGGAGTTTCAGACACAGCCTG
GCCAACTTGGTGAAAACCCCGTCTCTACTAAAAATACAAAAATTAGCTGGGCATGTTGGCAGG
CGCCTGTAATCCTTACCTACTTGGGAAGCTGAGGCAGGAGAATCACTTGAACCTGGGAGACGG
AGGTTGCAGTGAGCCAAAGATCACACCATTGCACGCCAGCCTGGGCAACAAGCGAGACTCCA
TCTCAAAAAAATCCTCCAATGGGTTGGGTGTCTGTAATCCCAGCACTTTGGGAGGCTA
AGGTGGGTGGATTGCTTGGAGCCAGGAGTTTCAGACACAGCCTGGGCAACATGTTGAAACCCC
ATCTCTACAAAAAATACAAAAATAGCTGGGCTGGGTGGTGTGTGCTGCTAGTGTCCAGTCT
CAGACATTTAAACCAGAGCAACTCCATCTGGAATAGGAGCTGGAATAAATAGAGGCTGAGGTC
TACTGGGCTGCATTCTCAGACAGTGGAGGCATTCTAAGTCACAGGATGAGACAGGAGGTCCG
TACAGATACAGGT CATAAAGACTTTGTCTGATAAAACAGATTGCAGTAAGAAGGCCAACCCCA
ATCCACCAAAACCAAGTTGGCCACAGGAGTGACCTCTGGCTGCTCACTGCTACACTCTCT
GACAGCACCATGACAGTTTACAAATGCCATGGCAACATCAGGAAGTTACCCGATATGTCCCA
AAAGGGGGGAAGTAATAATCCACCCCTTGTTTAGCAAATAAGCAAGAAATAACCATAAAA
GTGGCAACCAAGCTCTAGGCGCTGCTCTTGTCTATGGATAGCCATCTCTTTGTCTCTT
TACTTCTTAATAAACCTGCTTTTACCTTAAAAAAA

FIGURE 93

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA54002

><subunit 1 of 1, 544 aa, 1 stop

><MW: 60268, pI: 9.53, NX(S/T): 3

MLLPLLLSSLLGGSQAMDGRFWIRVQESVMVPEGLCISVPCSFSPRQDWTGSTPAYGYWFK
AVTETTTKAPVATNHQSREVEEMSTRGRFQLTGDPKAGNCSLVIRDAQMQDESQYFFRVERGS
YVTYNFMNDGFFLKVTLSFTPRPQDHNTDLTCHVDFSRKGVSAQRTVRLRVAYAPRDLVIS
ISRDNTPALEPQPQGNVPYLEAQKGQFLRLLLCAADSQPPATLSWVLQNRVLSSSHWPWGPRL
GLELPGVKAGDSGRYTCRAENRLGSQQRALDLSVQYPPENLRVMVSQANRTVLENLNGTSL
PVLEGQSLCLVCVTHSSPPARLSWTQRGQVLSPSQPSDPGVLELPRVQVEHEGEFTCHARHP
LGSQHVSLSLSVHYKKGLISTAFSNGAFLGIGITALLLFLCLALIIMKILPKRRTQTETPRPR
FSRHSTILDYINVVPTAGPLAQKRQKATPNSPRTPPPPGAPSPESKKNQKKQYQLPSFPEP
KSSTQAPESQESQEELHYATLNFPGVVRPRPEARMPKGTQADYAEVKFQ

Important features:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 399-418

N-glycosylation site.

amino acids 100-103, 297-300 and 306-309

Immunoglobulins and major histocompatibility complex proteins signature.

amino acids 365-371

FIGURE 94

TGAAGAGTAATAGTTGGAATCAAAAGAGTCAACGCAATGAACCTGTTATTTACTGCTGCGTTT
TATGTTGGGAATTCCTCTCCTATGGCCCTTGTCTTGGAGCAACAGAAAACTCTCAAACAAAGA
AAGTCAAGCAGCCAGTCGCGATCTCATTGTGAGAGTGAAGCTGGCTGGGTGTGGAAACCAATTT
TTTGTACAGAGGAAATGAATACGACTAGTGCATCACATCGGCCAGCTAAGATCTGATTTAGA
CAATGGAAACAATTCCTTCAGTACAAGCTTTTGGGAGCTGGAGCTGGAAGTACTTTTATCA
TTGATGAAAGAACAGGTGACATATATGCCATACAGAAGCTTGATAGAGAGGAGCGATCCCTC
TACATCTTAAGAGCCCAGGTAATAGACATCGCTACTGGAAGGGCTGTGGAACCTGAGTCTGA
GTTTGTCTCAAGAGTTTCGGATATCAATGACAATGAACCAAAATTCCTAGATGAACCTTATG
AGGCCATTGTACAGAGATGTCTCAGAAGGAACATTAGTTATCCAGGTGACAGCAAGTGAT
GCTGACGATCCCTCAAGTGGTAATATGCTCGTCTCCTCTACAGCTTACTTCAAGGCCAGCC
ATATTTTTCTGTTGAACCAACAACAGGAGTCATAAGAATATCTTCTAAAATGGATAGAGAAC
TGCAAGATGAGTATTGGGTAACTCATTCAGCCCAAGGACATGATTGGTCAGCCAGGAGCGTTG
TCTGGAACAACAAGTGATTAATTAACCTTTTCAGATGTTAATGACAATAAGCCTATATTTAA
AGAAAGTTTATACCGCTTGACTGTCTCTGAATCTGCACCCACTGGGACTTCTATAGGAACAA
TCATGGCATATGATAATGACATAGGAGAGAATGCAGAAATGGATTACAGCATTGAAGAGGAT
GATTCCGAAACATTTGACATTATTAATCAATGAAACTCAAGAGGGAATAGTTATATTTAA
AAAGAAAGTTGGATTTTGGAGCACAGAACCACACTACGGTATTAGAGCAAAAGTTAAAAACCATC
ATGTTCTCTGAGCAGCTCATGAAGTACCACACTGAGGCTTCCACCACCTTCATTAAAGATCCAG
GTGGAAGATGTTGATGAGCCTCCTCTTTCTCCTTCCATATTATGTATTTGAAGTTTGTGA
AGAAACCCCAAGGGATCATTTGTAGGCGTGGTGTCTGCCACAGACCCAGACCAATAGGAAAT
CTCCTATCAGGTATTTCTATTACTAGGAGCAAAGTGTTCAATATCAATGATATATGGTACAATC
ACTACAAGTATCTCAGTGACTCGTGAATCAGTGTCTTGGTACAACTAAGTATTACAGCCAC
AGAAAAATACAATATAGAACAGATCTCTTCGATCCCACTGTATGTGCAAGTTCTTAACATCA
ATGATCATGCTCCTGAGTTCTCTCAATACTATGAGACTTATGTTTGTGAAAATGCAAGGCTCT
GGTCAGGTAATTCAGACTATCAGTGCAGTGGATAGAGATGAATCCATAGAAGAGCACCATTT
TTACTTTAATCTATCTGTAGAAGACACTAACCAATTCAAGTTTTACAATCATAGATAATCAAG
ATAACACAGCTGTCAATTTTGACTAATAGAACTGGTTTTAACCTTCAAGAAGAACCTGTCTTC
TACATCTCCATCTTAATTGCCGACAATGGAATCCCGTCACTTACAAGTACAAACACCTTTAC
CATCCATGTCTGTGACTGTGGTGACAGTGGGAGCACACAGACCTGCCAGTACCAGGAGCTTG
TGCTTTCCATGGGATTCAGACAGAAGTTATCATTGCTATTCTCATTTTGCATTATGATCATA
TTTGGGTTTATTTTTTTGACTTTGGGTTTAAAAACAACGGAGAAAAACAGATTCTATTTCTG
GAAAAGTGAAGATTTAGAGAGAATATATTTCCAATATGATGATGAAGGGGGTGGAGAAGAG
ATACAGAGGCTTTGATATAGCAGAGCTGAGGAGTAGTACCATAATGCGGGAACGCAAGACT
CGGAAAACCAACAGCGCTAGATCAGGAGCCTATACAGGCAGTCTTTCGAAGTTGGCCCGCA
CAGTGCCATTTAGCAGAAATTCATTCTGGAAAAGCTCGAAGAAGCTAATCTGATCCGTGTG
CCCTCTCTTTTGATTCCTCCAGACCTACGCTTTTGAGGGAACAGGGTCATTAGCTGGATCC
CTGAGCTCTCTTGAATCAGCAGTCTCTGATCAGGATGAAAGCTATGATTACCTTAATGAGTT
GGGACCTCGCTTTAAAAGATTAGCATGCATGTTTGGTTCTGCGAGTCAATCAATTAATGAGG
GCTTTTTACCATCAAATTTTTAAAAGTGCTAATGTGTATTGCAACCCCAATGGTAGTCTTAA
AGAGTTTTGTGCCCCGTGCTATGGCGGGGAAAGCCCTAGTCTATGGAGTTTCTGATTTC
CTGAGTAATACTCCATGTTTATTTTAAGCTACCTACATGCTGTCAATGTAACAGAGATGTG
GGGAGAATGTAAACAATCAGCTCACAGGCATCAATACAACCAGATTTGAAGTAAATATATG
TAGGAAGATATTAAGTAGATGAGAGGACACAAGATGTAGTCGATCCTTATCGGATTATAT
CATTTATTTACTTAGAAAGAGTAAAAATACCAACGAGAAAAATTTAAAGGAGCAAAAATTTG
CAAGTCAAATAGAAATGTACAAATCGAGATAACATTACATTTCTATCATATTGACATGAAA
ATTGAAATGTATAGTCAGAGAAATTTTCATGAATTTATTCATGAAGTATTGTTTCCTTTAT
TTAA

FIGURE 95

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53906
><subunit 1 of 1, 772 aa, 1 stop
><MW: 87002, pI: 4.64, NX(S/T): 8
MNCYLLLRFMLGIPLLWPCLGATENSQTKVKQPVRSHLRVKRGWVWNQFFVPEEMNTTSHH
IGQLRSDLDNGNNSFQYKLLGAGAGSTFIIDERTGDIYAIQKLDREERSLYILRAQVIDIAT
GRAVEPESEFVIKVSDDINDNEPKFLDEPYEAIVPEMSPEGTLVIQVTASDADDPSSGNNARL
LYSLLQGQPYFSVEPTTGIVIRISSKMDRELQDEYVWIIQAKDMIGQPGALSGTTSVLIKLSL
VNDNKPIPKESLYRLTVSESAPTGTSGITIMAYDNDIGENAEMDYSIEEDDSQTFDIITNHE
TQEGIVILKKKVDFEHQNHYGIRAKVKNHHVPEQLMKYHTEASTTFIKIQVEDVDEPPLFLL
PYYVFEVFEETPQGSFVGVSATDPDNRKSPIRYSITRSKVFNINDNGTITTSNSLDREISA
WYNLSITATEKYNIEQISSIPLYVQVLNINDHAPEFSQYYETYVCENAGSGQVIQTISAVDR
DESBEEHHFYFNLSVEDTNNSSTFIIDNQDNTAVILTNRGTGNLQEEPVFYISILIADNGIP
SLTSTNTLTIHVDCDGSSTQTCQYQELVLSMGFKTEVIAILICIMIIFGFIFLTGLGLKQ
RRKQILFPKESDFRENIFQYDDEGGGEEDTEAFDIAELRSSTIMRERKTRKTTTABIRSLY
RQSLQVGPDSAIFRKFILEKLEEANTDPCAPPFDSLQTYAFEGTGSGLAGSLSSLESASVSDQD
ESYDYLNELGPRFKRLACMFGSAVQSNN

Important features:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 597-617

N-glycosylation sites.

amino acids 57-60, 74-77, 419-423, 437-440, 508-511, 515-518,
516-519 and 534-537

Cadherins extracellular repeated domain signature.

amino acids 136-146 and 244-254

FIGURE 96

ATTTCAAGGCCAGCCATATTTTNTGTTGAACCAACAACAGGAGTCATAAGAATATTTNTA
AAATGGATAGAGAACTGCAAGATGAGTATTGGGTAATCATTCAAGCCAAGGACATGATTGGT
CAGCCAGGAGCGTTGNTGGAACAACAAGTGTATTAATTAACTTTCAGATGTTAATGACAA
TAAGCCTATATTTAAAGAAAGTTTATACCGCTTGACTGTNTNTGAATCTGCACCCACTGGGA
NTTNTATAGGAACAATCATGGCATATGATAATGACATAGGAGAGAATGCAGAAATGGATTAC
AGCATTGAAGAGGATGATTGCGAAACATTGACATTATT

100
90
80
70
60
50
40
30
20
10
0
100
90
80
70
60
50
40
30
20
10
0

FIGURE 97

GCAACCTCAGCTTCTAGTATCCAGACTCCAGCGCGGCCCCGGGCGCGGACCCCCAACCCCGAC
CCAGAGCTTCTCCAGCGCGGCGCAGCGAGCAGGGCTCCCCGCCCTTAACCTTCTCCCGGGG
CCAGACCACTTCTCGGGAGTCCGGGTGCGCCACCTGCAAACTCTCCGCCCTTCTGCACCTGCCA
CCCCTGAGCCAGCGCGGGCCCCGAGCGAGTCAATGGCCAACGCGGGGCTGCAGCTGTTGGGC
TTCATTCTCGCCTTCTCGGGATGGATCGGCGCCATCGTCAGCAGCTGCCCTGCCCCAGTGGAG
GATTTACTCTTATGCCCGGCGACAACATCGTGACCGCCAGGCCATGTACGAGGGGCTGTGGA
TGTCTCGCTGTGCGCAGAGCACCGGGCAGATCCAGTGCAAAAGTCTTTGACTCCTTGCTGAAT
CTGAGCAGCACATGCGAAGCAACCCGTGCCTTGATGGTGGTTGGCATCCTCCTGGGAGTGAT
AGCAATCTTTGTGGCCACCGTTGGCATGAAGTGTATGAAGTGCTTGGAAAGCATGAGGTGC
AGAAGATGAGGATGGCTGTCAATGGGGGTGCGATATTTCTTCTTGAGGTCTGGCTATTTTA
GTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAATTTCTATGACCCCTATGACCCCAAGT
CAATGCCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCC
TTCTGGGAGGTGCCCTACTTTTGCTGTTCTGTGCCCGAAAAACAACCTCTTACCCAACACCA
AGGCCCTATCCAAAACCTGCACCTTCCAGCGGGAAGACTACGTGTGCACAGAGGCAAAAG
GAGAAAAATCATGTTGAACAAACCGAAAAATGGACATTTAGATACATCATTAACATTAGGAC
CTTAGAATTTTGGGTATTGTAATCTGAAGTATGGTATTACAAAAACAAACAAACAAAAA
ACCCATGTGTAAAATACTCAGTGTCTAAACATGGCTTAATCTTATTTTCTTCTTTCTCTCA
ATATAGGAGGGAAGATTTTCCATTTGTATTACTGCTTCCCATTTAGTAACTACACTCAAAAT
GGGGGAAGGGGTGCTCCTTAAATATATATAGATATGTATATATACATGTTTTTCTATTAAAA
ATAGACAGTAAAAATACTATTCTCATTATGTTGATAGCTACATCTAAAAATATCTCTAAAAAT
AGGTAATATGATTTAAATCCATATTGATGAAGATGTTTTATTGGTATATTTCTTTTCTGCTC
TTATATACATATGTAAACAGTCAAAATATCATTTTACTCTTCTTCACTTAGCTTGGGTGCCCTTTG
CCACAAGACCTAGCCTAATTTACCAAGGATGAATTTCTTCAATTTCTCATGGCTGCCCTTTT
CATATACTTATTTTATTTTTTACCATAATCTTATAGCACTTGCATCGTTATTAAAGCCCTTAT
TTGTTTTGTGTTTTATTGGTCTCTATCTCCTGAATCTAACACATTTATAGCCTACATTTTA
GTTTTCTAAAGCCAAGAAGAAATTTATTACAATCAGAAGTTTGGAGGCAAACTTTTCTGCATG
ACCAAAGTGATAAATTTCTGTTGACCTTCCCACACAATCCCTGTACTCTGACCCATAGCACT
CTTGTTTTGCTTTTGAAAAATATTTGCCAATTGAGTAGCTGCGATCTGTTCCCCAGGGTGTGT
AACACAACCTTTATTTGATTGAATTTTTTAAGCTACTTATTCATAGTTTTATATCCCCCTAAACT
ACCTTTTTTGTTCCTCCATCTTAATTTGATTGTTTTTCCCAAAGTGAATTTATCATGCGTTTTA
TATCTTCTAATAAGGTGTGGTCTGTTTTGTCTGAACAAAGTGCTAGACTTTCTGGAGTGATA
ATCTGGTGACAAATATTTCTCTGTAGCTGTAAGCAAGTCACTTAATCTTTCTACCTCTTTT
TTCTATCTGCCAAATTGAGATAATGATACTTAACCAAGTLAGAAGAGGTAGTGTAATATTAA
TTAGTTTATATTACTCTTATCTTTGAACATGAACATAATGCTATGTAGTGTCTTTATTGGCT
CAGCTGGCTGAGACACTGAAGAAGTCACTGAACAAAAACCTACACACGTACCTTCATGTGATT
CACTGCCCTTCCCTCTCTACCAAGTCTATTTCCACTGAACAAAACCTACACACATACCTTCAT
GTGGTTTCAGTGCCCTCTCTCTCTACCAAGTCTATTTCCACTGAACAAAACCTACGCACATAC
CTTCATGTTGGCTCAGTGGCTTCCCTCTCTCTACCAAGTCTATTTCCATTTCTTCCAGTGTGTCT
GACATGTTTTGTGCTCTGTTCCATTTTAAACAAGTCTCTTACTTTTCCAGTCTGTACAGAATG
CTATTTCACTTGAGCAAGATGATGTAATGGAAGGGGTGTTGGCACTGGTGTCTGGAGACCTG
GATTTGAGTCTTTGGTGCTATCAATCACCGTCTGTGTTTGAGCAAGGCATTTGGCTGCTGTAA
GCTTATTGCTTCATCTGTAAGCGGTGGTTTGTAAATTCCTGATCTTCCCACCTCACAGTGATG
TTGTGGGGATCCAGTGAGATAGAATACATGTAAGTGTGGTTTTGTAAATTTAAAAAGTGCTAT
ACTAAGGGAAAGAAATGAGGAATTAACCTGCATACGTTTTGGTGTGTTTTGCTTTTCAAAATGTTGA
AAATAAAAAAATGTTAAG

FIGURE 98

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52185
><subunit 1 of 1, 211 aa, 1 stop
><MW: 22744, pI: 8.51, NX(S/T): 1
MANAGLQLLGFI~~L~~AWIGAI~~V~~STALPQWRIYSYAGDNIVTAQAMYEGLWMSCV~~S~~QSTGQI
QCKVFD~~S~~LLNLSSTLQATRALMVVGILLGVIAIFVATVGMKCMKCLEDE~~V~~QKMRMAVIGGA
IFLLAGLAILVATAWYGNRIVQEFYDPMTFVNARYEFGQALFTGWAAASLCLLG~~G~~ALLCCSC
PRK~~T~~TSYPTPRPYPKPAPSSGKDYV

Important features:

Signal peptide:

amino acids 1-21

Transmembrane domains:

amino acids 82-102, 118-142 and 161-187

N-glycosylation site.

amino acids 72-75

PMP-22 / EMP / MP20 family proteins

amino acids 70-111

ABC-2 type transport system integral membrane protein

amino acids 119-133

FIGURE 99

TTCTGGCCAAACCCGGGGCTNCAGCTGTTGGGCTTCATCTCGCCTTCCTGGGATGGATCGGC
GCCATCNTCACACTGCCCTTCCCCAGTGGAGGATTTTACTCCCTATGCTGGCGACAACATCG
TGACCGCCCCAGCCCATGTACGAGGGGCTGTGGATGTCCNGCGTGTGCGAGAGCACCGGGCAG
ATCCAGTGCAAAGTCTTTGACTCCTTGCTGAATCTGAGCAGCACATTGCAAGCAACCCGTGC
CTTGATGGTGTTGGCATCCTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTTGGCATGA
AGTGTATGAAGTGCTTGAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCATTGGGGGC
GCGATATTTCTTCTTGCAAGTCTGGCTATTTTAGTTGCCACAGCATGGTATGGCAATAGAA
CNTTCAACANTTCTATGACCCATGACCCAGTCAATGCCAGGTACGAATTTGGTCA
GGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCCTTCTGGGAGGTGCCCTACTTTGCT
GTTCCTGTCCC

100
90
80
70
60
50
40
30
20
10
0

FIGURE 100

ACCCTTGACCCAACGCGGCCCCCGACCGNTTCATGGCCAAACGCGGGNCTCCAGCTGTTGG
GCTTCATTCTCCCTTCTCTGGGATGGACCGGCGCCCATCNTCAGCACTGCCCTGCCCCAGTG
GAGGATTTACTCCTATNCCGGCNACAACATCGTGACCGCCCAGGCCNTGTACGAGGGGCTGT
GGATGTCTGCGTGTGCGCAGAGCACCGGGCAGATCCAGTGCAAAGTCTTTGACTCCCTTGCT
GAATCTGAGCAGCACATTGCAAGCAACCCGTGCCTTGATGGTGGTTGGCATCCTCCTGGGAG
TGATAGCAATCTTNNTGGCCACCGTTGTNNNTGAAGTGATGAAGTGCTTGGAGACGATGA
GGTGCAGAAGATGAGGATGGCTGTTCATTGGGGGCGCGATATTTCTTCTTGCAAGTCTGGCTA
TTTtagTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAATTCTATGACCCATGACCGA

100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200

FIGURE 101

GGGCCCACCATTATCCAACCGGGNTCACTGTTGGCTCATCTCCCTCCTGGATGAANCGCGC
CATCNTCAGACTCCCTGCCCCATGGAGATTTNNCCTATGCTGGCGACAACATCNTGACCCCC
AGCCATGTACGAGGGGCTTTGAACGTCNGCGTGTGCGAGANACCGGGCAGATCCAGTGCAA
AGTCTTTGACTCCTTGCTGAATCTGNGCAGCACATTGCAGCAACCCNTGCCCTGATGGTGGT
TGGCATCCTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTTGGCATGAAGTGATGAAGT
GCTTGGAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCATTGGGGGCGCGATATTTCTT
CTTGCAAGTCTGGCTATTTNNNGTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAAT
TCTATGACCCATGACCCAGTCAATGCCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGC
TGGGCTGCTGCTTCTCTCTGCCTTCTGGGAGGTGCCCTACTTTGCTGTTCTCTGCGA

63
143
243
343
443
543
643
743
843
943
1043
1143
1243
1343
1443
1543
1643
1743
1843
1943
2043
2143
2243
2343
2443
2543
2643
2743
2843
2943
3043
3143
3243
3343
3443
3543
3643
3743
3843
3943
4043
4143
4243
4343
4443
4543
4643
4743
4843
4943
5043
5143
5243
5343
5443
5543
5643
5743
5843
5943
6043
6143
6243
6343
6443
6543
6643
6743
6843
6943
7043
7143
7243
7343
7443
7543
7643
7743
7843
7943
8043
8143
8243
8343
8443
8543
8643
8743
8843
8943
9043
9143
9243
9343
9443
9543
9643
9743
9843
9943
10043

1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 20. 21. 22. 23. 24. 25. 26. 27. 28. 29. 30. 31. 32. 33. 34. 35. 36. 37. 38. 39. 40. 41. 42. 43. 44. 45. 46. 47. 48. 49. 50. 51. 52. 53. 54. 55. 56. 57. 58. 59. 60. 61. 62. 63. 64. 65. 66. 67. 68. 69. 70. 71. 72. 73. 74. 75. 76. 77. 78. 79. 80. 81. 82. 83. 84. 85. 86. 87. 88. 89. 90. 91. 92. 93. 94. 95. 96. 97. 98. 99. 100. 101. 102. 103. 104. 105. 106. 107. 108. 109. 110. 111. 112. 113. 114. 115. 116. 117. 118. 119. 120. 121. 122. 123. 124. 125. 126. 127. 128. 129. 130. 131. 132. 133. 134. 135. 136. 137. 138. 139. 140. 141. 142. 143. 144. 145. 146. 147. 148. 149. 150. 151. 152. 153. 154. 155. 156. 157. 158. 159. 160. 161. 162. 163. 164. 165. 166. 167. 168. 169. 170. 171. 172. 173. 174. 175. 176. 177. 178. 179. 180. 181. 182. 183. 184. 185. 186. 187. 188. 189. 190. 191. 192. 193. 194. 195. 196. 197. 198. 199. 200. 201. 202. 203. 204. 205. 206. 207. 208. 209. 210. 211. 212. 213. 214. 215. 216. 217. 218. 219. 220. 221. 222. 223. 224. 225. 226. 227. 228. 229. 230. 231. 232. 233. 234. 235. 236. 237. 238. 239. 240. 241. 242. 243. 244. 245. 246. 247. 248. 249. 250. 251. 252. 253. 254. 255. 256. 257. 258. 259. 260. 261. 262. 263. 264. 265. 266. 267. 268. 269. 270. 271. 272. 273. 274. 275. 276. 277. 278. 279. 280. 281. 282. 283. 284. 285. 286. 287. 288. 289. 290. 291. 292. 293. 294. 295. 296. 297. 298. 299. 300. 301. 302. 303. 304. 305. 306. 307. 308. 309. 310. 311. 312. 313. 314. 315. 316. 317. 318. 319. 320. 321. 322. 323. 324. 325. 326. 327. 328. 329. 330. 331. 332. 333. 334. 335. 336. 337. 338. 339. 340. 341. 342. 343. 344. 345. 346. 347. 348. 349. 350. 351. 352. 353. 354. 355. 356. 357. 358. 359. 360. 361. 362. 363. 364. 365. 366. 367. 368. 369. 370. 371. 372. 373. 374. 375. 376. 377. 378. 379. 380. 381. 382. 383. 384. 385. 386. 387. 388. 389. 390. 391. 392. 393. 394. 395. 396. 397. 398. 399. 400. 401. 402. 403. 404. 405. 406. 407. 408. 409. 410. 411. 412. 413. 414. 415. 416. 417. 418. 419. 420. 421. 422. 423. 424. 425. 426. 427. 428. 429. 430. 431. 432. 433. 434. 435. 436. 437. 438. 439. 440. 441. 442. 443. 444. 445. 446. 447. 448. 449. 450. 451. 452. 453. 454. 455. 456. 457. 458. 459. 460. 461. 462. 463. 464. 465. 466. 467. 468. 469. 470. 471. 472. 473. 474. 475. 476. 477. 478. 479. 480. 481. 482. 483. 484. 485. 486. 487. 488. 489. 490. 491. 492. 493. 494. 495. 496. 497. 498. 499. 500. 501. 502. 503. 504. 505. 506. 507. 508. 509. 510. 511. 512. 513. 514. 515. 516. 517. 518. 519. 520. 521. 522. 523. 524. 525. 526. 527. 528. 529. 530. 531. 532. 533. 534. 535. 536. 537. 538. 539. 540. 541. 542. 543. 544. 545. 546. 547. 548. 549. 550. 551. 552. 553. 554. 555. 556. 557. 558. 559. 560. 561. 562. 563. 564. 565. 566. 567. 568. 569. 570. 571. 572. 573. 574. 575. 576. 577. 578. 579. 580. 581. 582. 583. 584. 585. 586. 587. 588. 589. 590. 591. 592. 593. 594. 595. 596. 597. 598. 599. 600. 601. 602. 603. 604. 605. 606. 607. 608. 609. 610. 611. 612. 613. 614. 615. 616. 617. 618. 619. 620. 621. 622. 623. 624. 625. 626. 627. 628. 629. 630. 631. 632. 633. 634. 635. 636. 637. 638. 639. 640. 641. 642. 643. 644. 645. 646. 647. 648. 649. 650. 651. 652. 653. 654. 655. 656. 657. 658. 659. 660. 661. 662. 663. 664. 665. 666. 667. 668. 669. 670. 671. 672. 673. 674. 675. 676. 677. 678. 679. 680. 681. 682. 683. 684. 685. 686. 687. 688. 689. 690. 691. 692. 693. 694. 695. 696. 697. 698. 699. 700. 701. 702. 703. 704. 705. 706. 707. 708. 709. 710. 711. 712. 713. 714. 715. 716. 717. 718. 719. 720. 721. 722. 723. 724. 725. 726. 727. 728. 729. 730. 731. 732. 733. 734. 735. 736. 737. 738. 739. 740. 741. 742. 743. 744. 745. 746. 747. 748. 749. 750. 751. 752. 753. 754. 755. 756. 757. 758. 759. 760. 761. 762. 763. 764. 765. 766. 767. 768. 769. 770. 771. 772. 773. 774. 775. 776. 777. 778. 779. 780. 781. 782. 783. 784. 785. 786. 787. 788. 789. 790. 791. 792. 793. 794. 795. 796. 797. 798. 799. 800. 801. 802. 803. 804. 805. 806. 807. 808. 809. 810. 811. 812. 813. 814. 815. 816. 817. 818. 819. 820. 821. 822. 823. 824. 825. 826. 827. 828. 829. 830. 831. 832. 833. 834. 835. 836. 837. 838. 839. 840. 84

ATTCTCCCTCTCTGGATGGATCGCNCACCGTCACATTGCCTTCCCCCANTGGAGGATTNAC
TCCTATGCTGGCGACAACATCGTGACCCCCAGGCCATTACCGAGGGGCTTTGGATGTCNT
GCNTGTGCGAGAGCACCGGGCAGATCCCAGTGCAAAGTCTTTGACTCCTTGCTGAATCTGAG
CAGCACATTGCAAGCAACCCGTGCCTTGATGGGGTTGGCATCTCTCTGGGAGTGATAGCAAC
CTTTGTGGCCACCGTTGGCATGAAGTGATGAAGTGCTTGAAGACGATGAGGTGCCAGAAG
ATGAGGATGGCTGTCAATTGGGGGCGCGATATTTCTTGTTGCAGGCTCTGGCTATTTTAGTNGC
CACAGCATGGTATGGC AATAGANTNNTT CNNGNNTCTATGACCCCTATGACCCAGTCAATG
CAGGATACGAATTTGGCTAGGCTCTTCTTCACTGGCTGGGCTGCTGCTTCTCTGCTCTCTG
GGAGGTGCCCTACTTTTGCTTCTCTGCC

Figure 1. The structure of the proposed model.

AGAGCACCGGCAGATCCCAGTNCAAAGTCTTTGACCCCTTGCTGAATCTGAGCAGCACATTNC
AAGCAACCCCTTGCTTGAAGGTGGTTGNCATCCCCCTGGGAGTGAAATAGCAATCTTTGTG
GCCACCGTTGGCATGAAGTNTATGAAGTGCTTGGGAAGACGATGAGGTGCAGAAGATGAGGAT
GGCTGTCAATTGGGGGCGCATATTTCTTCTTGACAGGTCTGGCTATTTTAGTNNCCACAGCAT
GGTATGGCAATAGNATNNTTGNGNNNTTCTATGACCCCTATGACCCCAAGTCAATGCCAGGTAC
GAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCTTCTGGGAGGTGC
CCTACTTTGCTGTTCTCTGCTCCCGGAA

FIGURE 104

AGCAATGCCCTGCCCCAGTGGAGGATTAATTCCCTATGNTGGGGACAACATTGTGACNGCCC
AGGCCATGTACGGGGGGCTGTGGATGTCCTGCGTGTGCGAGAGCACCGGGCAGATCCAGTGC
AAAGTNTTTGACTCCTTGCTGAATTTGAGCAGCACATTGCAAGCAACCCGTGCCTTGATGGT
GGTTGGCATCTTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTGGNAATGAAGTGATGA
AGTGCTTGGAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCATTGGGGGCGCGATATTT
CTTNTTGCAGGTCTGGCTATTTTAGTTGCCACAGCATGGTATGGCAATAGAATNGTTCAAGA
ATTTTATGACCCATGACCCAGTCAATGCCAGGTACGAATTTGGTCAGGCTTTNTTCACTG
GCTGGGCTGCTGCTTNTTCTGCCTTNITGGGAGGTGCCCTANTTTGCTGTTCTCGCAACC

63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 106

TTCTGCGGATGGATCCGCCCCCATCNTCACATGCCCTGCCCCNTGGAGATTTACNCCTATGC
TGGCGAACAACATCNTGACCGCCCAGGCCATGTACGAGGGGCTGTGGAATGTCCTGCGTGTC
CCAGAGCACCGGGCAGATCCAGTGCAAAGTCTTTGACTCCTTGCTGAATCTGAGCAGCACAT
TGCAAGCAACCNTGCCTTGATGGTGGTTGGCATCCTCCTGGGAGTGATAGCAATCTTTGTGG
CCACCGTTGGCATGAAAGTGATGAAGTGCTTGAAGACGATGAGGTGCAGAAGATGAGGAT
GGCTGTCATTGGGGGCGCGATATTTCTTCTTGCAGGTCTGGCTATTTTAGNNGCCACAGCAT
GGTATGGCAATCAGACCCNNTCANAACTCTATGACCCTATGACCCAGTCAATGCCAGGTA
CGAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTGCTTCTGGGAGGTG
CCCTACTTTGCTGTTCTGTCCCCGAAAAACAACCTCTTACCCACG

100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212
2213
2214
2215
2216
2217
2218
2219
2220
2221
2222
2223
2224
2225
2226
2227
2228
2229
2230
2231
2232
2233
2234
2235
2236
2237
2238
2239
2240
2241
2242
2243
2244
2245
2246

FIGURE 107

CGGGGCTGCAGCTGTTGGGCTTCATCTCGCTTCCTGGGATGGAATCGGCGCCATCGTCAGCA
CTGCCCTGCCCCATGGAGGATTTACTCNTATGCTGGCGACAACATCGTGACCNCCCAGGCCA
TGTACGAGGGGCTGTGGATGTNGCGTGTGCGAGAGCACCGGGCAGATCCAGTGCAAAGTCT
TTGACTCCTTGCTGAATCTGAGCAGCACATTGCAAGCAACCNCGCCTTGATGGTGGTTGGCA
TCCTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTTGGCATGAAGTGATGAAGTGCTTG
GAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCATTGGGGGCGCGATATTTCTTCTTGC
AGGTCTGGCTATTTNTAGTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAATCTAT
GACCCTATGACCCAGTCAATGCCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGCTGGGC
TGCTGCTTCTCTGCTTCCTGGGAGGTGCCCTACTTTGCTGTTCTGCGAA

100
90
80
70
60
50
40
30
20
10
0
100
90
80
70
60
50
40
30
20
10
0

FIGURE 108

GCGTGCCGTGAGCTCGCCGGGACACCGCGGCTCGCCCTCGCCCTCCGCCCTGCGCCTGCAC
CGCGTAGACCGACCCCCCTCCAGCGCGCCACCCGGTAGAGGACCCCGCCCGTGCCCCG
ACCGGTCGCCCTTTTTGTAAACTTAAAGCGGGCGCAGCATTAAACGTTCCCGCCCCGT
GACCTCTCAGGGGTCTCCCGCCAAAGGTGCTCCGCGCTAAGGAACATGGCGGAAGGTGGAG
CAGGTCTGAGCCTCGAGCCGCGAGCAGAGCTCAAATTCGAGGTCCCTTACCCGATGTTGT
CACCACCAACCTAAAGCTTGGCAACCCGACAGACCGAAATGTGTGTTTTAAGGTGAAGACTA
CAGCACCACGTAGGTACTGTGTGAGGCCCAACAGCGGAATCATCGATGCGAGGGCCTCAATT
AATGTATCTGTGATGTTACAGCCTTTCGATTATGATCCCAATGAGAAAAAGTAAACACAAGTT
TATGGTTCAGTCTATGTTTGTCCAACTGACACTTCAGATATGGAAGCAGTATGGAAGGAGG
CAAAACCGGAAGACCTTATGGATTCAAACTTAGATGTGTGTTGAATTGCCAGCAGAGAAT
GATAAACCATGATGTAGAATAAATAAAATATATCCACAACCTGCATCAAAGACAGAAAC
ACCAATAGTGTCTAAGTCTCTGAGTTCCTCTTTGGATGACACCGAAGTTAAGAGGTTATGG
AAGAATGTAAGAGGCTGCAAGGTGAAGTTCAGAGGCTACGGGAGGAGAACCAAGCAGTTCAG
GAAGAAGATGGACTGCGGATGAGGAAGACAGTGCAGAGCAACAGCCCCATTTCCAGCATTAGC
CCCAACTGGGAAGGAAGAAGGCCTTAGCACCCGGCTCTTGCTCTGGTGGTTTTGTTCTTTA
TCGTTGGTGTAATTATTGGGAAGATTGCCCTTGTAAGGTAGCATGCACAGGATGGTAAATTG
GATTGGTGGATCCACCATATCATGGGATTTAAATTTATCATAACCATGTGTAAAAAGAAAT
AATGTATGATGACATCTCACAGGTCTTGCCTTTAAATTACCCCTCCCTGCACACACATACAC
AGATACACACACAAATATAATGTAACGATCTTTTGAAGAGTTAAAAATGTATAGTAACTG
ATTGAGGGGAAAAAGAATGATCTTTATTAATGACAAGGGAAACCATGAGTAATGCCACAAT
GGCATATTGTAATGTCAATTTTAAACATTGGTAGGCCTTGTGTACATGATGCTGGATTACCTC
TCTTAAATGACACCCCTCTTCGCTGTTGGTGCTGGCCCTTGGGAGCTGAGGCCCCAGCAT
GCTGGGGAGTGCGGTGAGCTCCACACAGTAGTCCCACGTGGCCACTCCCGGCCAGGCTG
CTTTCCTGTCTTCAGTTCGTCCAGCCATCAGCTCCCTGGGACTGATGAACAGAGTCAGA
AGCCCAAAGGAATTGCACTGTGGCAGCATCAGACGTACTCGTCATAAGTGAGAGGCGTGTGT
TGACTGATTGACCCGCGCTTGGAAATAAATGCGAGTGCCTTTGTTCACCTTAAAGGGACCAA
GCTAAATTTGATTTGGTTCATGTAGTGAAGTCAAACCTGTTATTACAGAGATGTTTAAATGCATA
TTTAACTTATTTAATGTATTTCACTCATGTGTTTTCTTATTGTACAAAGAGTACAGTTAATGC
TGCGTGCTGCTGAACTCTGTTGGGTGAACCTGGTATTGCTGCTGGAGGGCTGTGGGCTCCTCT
GTCTCTGGAGAGTCTGTCATGTGGAGGTGGGTTTTATTGGGATGCTGGAGAAGAGACTGCCA
GGAAGTGTTTTTCTGGGTGAGTAAATAACAACCTGTATAGGAGGGAAATCTCAGTAGTG
ACAGTCAACTCTAGGTTACCTTTTTAATGAAGAGTAGTCAGTCTCTCAGATTGTTCTTATA
CCACCTCTCAACCATTACTCACACTTCCAGCGCCAGGTCCAAGTCTGAGCCTGACCTCCCC
TTGGGGACCTAGCCTGGAGTCAGGACAAATGGATCGGGCTGCAGAGGGTTAGAAGCGAGGGC
ACCGCAGTGTGTGGTGGGAGCAAGGGAAGAGAACTCTTCAGCGAATCCTTCTAGTAC
TAGTTGAGAGTTTGAAGTGTGAATTAATTTTATGCCATAAAAGACCAACCCAGTTCTGTGTTGA
CTATGTAGCATCTGAAAAGAAAAATTATAATAAAGCCCCAAATTAAGAAAA

FIGURE 109

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53977

<subunit 1 of 1, 243 aa, 1 stop

<MW: 27228, pI: 7.43, NX(S/T): 2

MAKVEQVLSLEPQHELKFRGPFPTDVVTTNLKLGNPTRNVCFKVKTTAPRRYCVRPNSGIID
AGASINVSVMQLQPFDDPNEKSKHKFMVQSMFAPTDTSMEAVWKEAKPEDLMSKLRVFE
LPAENDKPHDVEINKIISTTASKTETPIVSKSLSSSLDDTEVKKVMEECKRLQGEVQRLREE
NKQFKEEDGLMRKTVQSNSPISALAPTGKEEGLSTRLLALVVLFFIVGVIIGKIAL

Important features:

Transmembrane domain:

amino acids 224-239

N-glycosylation site.

amino acids 68-71

N-myristoylation site.

amino acids 59-64, 64-69 and 235-240

FIGURE 110

GTCACTCTTCTAGATTGTCCTTATCCCACCTTTCAACCANTACTCACATTTTCNAGCGCCCAG
GTCCANGTCTGAGCCTGACTTCCCCTTGGGGACCTAGCCTGGAGTCAGGACAAATGGNTCGGG
CTGCAGAGGNTTAGAAGCGAGGGCACCAGCAGTTTTGGGTGGGGAGCAAGGGNNGAGAGAAA
CTCTTCAGCGAATCCTTCTAGTACTAGTTGAGAGTTTGACTGTGAATTAATTTTATGCCATA
AAAGACNAACCCAGTTCTGTTTGACTATGTAGCATCTGAAAAGAAAAATTATAATAAAGCC
CCAAAATTAAGAATTCTTTTGTCATTTTGTACATTTGCTCTATGGGGGAATTATTATTTT
ATCATTTTATTATTTTGCCATTGGAAGGTTAACTTTAAATGAGC

100
90
80
70
60
50
40
30
20
10
0
A
T
C
G
A
T
C
T
C
T
A
G
A
T
T
G
T
C
C
T
T
A
T
C
C
C
A
C
C
T
T
T
C
A
A
C
C
A
N
T
A
C
T
C
A
C
A
T
T
T
T
C
N
A
G
C
G
C
C
C
A
G
T
C
C
A
N
G
T
C
T
G
A
G
C
C
T
G
A
C
T
T
C
C
C
C
T
T
G
G
G
G
A
C
C
T
A
G
C
C
T
G
G
A
G
T
C
A
G
G
A
C
A
A
T
G
G
N
T
C
G
G
G
C
T
G
C
A
G
A
G
G
N
T
T
A
G
A
A
G
C
G
A
G
G
G
C
A
C
C
A
G
C
A
G
T
T
T
T
G
G
G
T
G
G
G
G
A
G
C
A
A
G
G
G
N
N
G
A
G
A
G
A
A
A
C
T
C
T
T
C
A
G
C
G
A
A
T
C
C
T
T
C
T
A
G
T
A
C
T
A
G
T
T
G
A
G
A
G
T
T
T
G
A
C
T
G
T
G
A
A
T
T
A
A
T
T
T
T
A
T
A
T
A
A
A
A
G
C
C
C
A
A
A
A
T
T
A
A
G
A
A
T
T
C
T
T
T
T
G
T
C
A
T
T
T
T
G
T
C
A
C
A
T
T
T
G
C
T
C
T
A
T
G
G
G
G
G
A
A
T
T
A
T
T
A
T
T
T
A
T
C
A
T
T
T
T
T
A
T
T
T
T
G
C
C
A
T
T
G
G
A
A
G
G
T
T
A
A
C
T
T
T
A
A
A
T
G
A
G
C

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|--------|
| 1990 | 1991 | 1992 | 1993 | 1994 | 1995 | 1996 | 1997 | 1998 | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 | 2020 | 2021 | 2022 | 2023 | 2024 | 2025 | 2026 | 2027 | 2028 | 2029 | 2030 | 2031 | 2032 | 2033 | 2034 | 2035 | 2036 | 2037 | 2038 | 2039 | 2040 | 2041 | 2042 | 2043 | 2044 | 2045 | 2046 | 2047 | 2048 | 2049 | 2050 | 2051 | 2052 | 2053 | 2054 | 2055 | 2056 | 2057 | 2058 | 2059 | 2060 | 2061 | 2062 | 2063 | 2064 | 2065 | 2066 | 2067 | 2068 | 2069 | 2070 | 2071 | 2072 | 2073 | 2074 | 2075 | 2076 | 2077 | 2078 | 2079 | 2080 | 2081 | 2082 | 2083 | 2084 | 2085 | 2086 | 2087 | 2088 | 2089 | 2090 | 2091 | 2092 | 2093 | 2094 | 2095 | 2096 | 2097 | 2098 | 2099 | 2100 | 2101 | 2102 | 2103 | 2104 | 2105 | 2106 | 2107 | 2108 | 2109 | 2110 | 2111 | 2112 | 2113 | 2114 | 2115 | 2116 | 2117 | 2118 | 2119 | 2120 | 2121 | 2122 | 2123 | 2124 | 2125 | 2126 | 2127 | 2128 | 2129 | 2130 | 2131 | 2132 | 2133 | 2134 | 2135 | 2136 | 2137 | 2138 | 2139 | 2140 | 2141 | 2142 | 2143 | 2144 | 2145 | 2146 | 2147 | 2148 | 2149 | 2150 | 2151 | 2152 | 2153 | 2154 | 2155 | 2156 | 2157 | 2158 | 2159 | 2160 | 2161 | 2162 | 2163 | 2164 | 2165 | 2166 | 2167 | 2168 | 2169 | 2170 | 2171 | 2172 | 2173 | 2174 | 2175 | 2176 | 2177 | 2178 | 2179 | 2180 | 2181 | 2182 | 2183 | 2184 | 2185 | 2186 | 2187 | 2188 | 2189 | 2190 | 2191 | 2192 | 2193 | 2194 | 2195 | 2196 | 2197 | 2198 | 2199 | 2200 | 2201 | 2202 | 2203 | 2204 | 2205 | 2206 | 2207 | 2208 | 2209 | 2210 | 2211 | 2212 | 2213 | 2214 | 2215 | 2216 | 2217 | 2218 | 2219 | 2220 | 2221 | 2222 | 2223 | 2224 | 2225 | 2226 | 2227 | 2228 | 2229 | 2230 | 2231 | 2232 | 2233 | 2234 | 2235 | 2236 | 2237 | 2238 | 2239 | 2240 | 2241 | 2242 | 2243 | 2244 | 2245 | 2246 | 2247 | 2248 | 2249 | 2250 | 2251 | 2252 | 2253 | 2254 | 2255 | 2256 | 2257 | 2258 | 2259 | 2260 | 2261 | 2262 | 2263 | 2264 | 2265 | 2266 | 2267 | 2268 | 2269 | 2270 | 2271 | 2272 | 2273 | 2274 | 2275 | 2276 | 2277 | 2278 | 2279 | 2280 | 2281 | 2282 | 2283 | 2284 | 2285 | 2286 | 2287 | 2288 | 2289 | 2290 | 2291 | 2292 | 2293 | 2294 | 2295 | 2296 | 2297 | 2298 | 2299 | 2300 | 2301 | 2302 | 2303 | 2304 | 2305 | 2306 | 2307 | 2308 | 2309 | 2310 | 2311 | 2312 | 2313 | 2314 | 2315 | 2316 | 2317 | 2318 | 2319 | 2320 | 2321 | 2322 | 2323 | 2324 | 2325 | 2326 | 2327 | 2328 | 2329 | 2330 | 2331 | 2332 | 2333 | 2334 | 2335 | 2336 | 2337 | 2338 | 2339 | 2340 | 2341 | 2342 | 2343 | 2344 | 2345 | 2346 | 2347 | 2348 | 2349 | 2350 | 2351 | 2352 | 2353 | 2354 | 2355 | 2356 | 2357 | 2358 | 2359 | 2360 | 2361 | 2362 | 2363 | 2364 | 2365 | 2366 | 2367 | 2368 | 2369 | 2370 | 2371 | 2372 | 2373 | 2374 | 2375 | 2376 | 2377 | 2378 | 2379 | 2380 | 2381 | 2382 | 2383 | 2384 | 2385 | 2386 | 2387 | 2388 | 2389 | 2390 | 2391 | 2392 | 2393 | 2394 | 2395 | 2396 | 2397 | 2398</ |
|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|--------|

TATTGTAAAGGCCATTTTAAACCATTGGTAGGCCTTGGTACATGATGCTGGATTACCTCCTT
AAATGACACCNTTCTCTCGCCTGTTGGTGCTGGCCNTTGGGGAGCTGGAGCCCCAGCATGCTG
GGGAGTGCGGTGAGCTCCACACAGTAGTCCCCACGTGGCCCACTCCGGGCCAGGCTGCTTT
CCGTGCTTTCAGTTCTGTCCAAGCCATCAGTCTCTTGGGACTGATGAACAGAGTCAGAAGCC
CAAAGGAATTGCCACTGTGGCAGCATCAGACGTACTCGTCATAAGTGAGAGGCGTGTGTGA
CTGATTGACCCAGCGCTTTGGAAATAAATGGCAGTGCTTTGTTCACTTAAAGGGACCAAGCT
AAATTGTATTGGTTTCATGTAGTGAAGTCAAACCTGTTATTAGAGATGTTTAAATGCATATTTA
ACTTATTTAATGTATTTTCATCTCATGTTTTCTTATTGTCAAGAGTACAGTTAATGCTGCG
TGCTGCTGAACCTCTGTTGGGTGAACCTGGTATTGCTGCTGGAGGGCTG

FIGURE 112

CCCTGGTGGTTTTGTTCTTTAATTCGTTGGTGTAATTNTTGGGAAGATTGCTTGTAGAGGTA
GNATGCACCNGGCTGGTAAATTGGATTGGTGGATCCACCATATCCATGGGATTTAAATTTAT
CATAACCATGTGTAAAAAGAAATTAATGTATGATGACATNTCACAGGTATTGCCTTTAAATT
ACCCATCCCTGNANACACATACAGATACACANANACAAATNTAATGTAACGATNTTTTAG
AAAGTTAAAAATGTATAGTAAC

100
90
80
70
60
50
40
30
20
10
0
100
90
80
70
60
50
40
30
20
10
0

FIGURE 113

GGTGGCCCATTCCTGGCCAGGCTGCTTTCCGGTNTTCAGTTCTGTCCAAGCCATCAGCTCC
TTGGGACTGATGAACAGAGTCAGAAGCCCAAAGGAATTGCACTGTGGCAGCATNAGACGTAC
TTGTNATAAGTGAGAGGCGTGTGTTGACTGATTGACCCAGCGCTTTGGAAATAAATGGCAGT
GCTTTGTTTCANTTAAAGGGACCAAGCTAAATTTGTATTGGTTCATGTAGTGAAGTCAAACG
TTATTCAGAGATGTTAATGCATATTTAANTTATTTAATGTATTTNATNTCATGTTTTCTTA
TTGTCACAAGAGTACAGTTAATGCTGCGTGCTGCTGAANTNTGTTGGGTGAAGTGGTATTGC
TGCTGGAGGGCTGTGGGCTCCTCTGCTTTGGAGAGTCTGGTCATGTGGAGGTGGG

1000
900
800
700
600
500
400
300
200
100
0
1000
900
800
700
600
500
400
300
200
100
0
1000
900
800
700
600
500
400
300
200
100
0

FIGURE 114

TGCTTTCGTTGTTCTTCAGTTCTGTCCAAGCCATCAGCTCCTTGGGACTTGATGAACAGAGTC
AGAAGCCCAAAGGAATTGCACTGTGGCAGCATCAGACGTACTCGTCATAAGTGAGAGGCGTG
TGTTGACTGATTGACCCAGCGCTTTGGAAATAAATGGCAGTGCTTTGTTCACTTAAAGGGAC
CAAGCTAAATTTGTATTGGTTCATGTAGTGAAGTCAAAGTGTATTTCAGAGATGTTTAAATGC
ATATTTAACTTATTTAATGTATTTTCATCTCATGTTTTCTTATTGTCACAAGAGTACAGTTAA
TGCTGCGTGC

CONCLUSIONS

A A A C C T T T A A A A G T T G A G G G G A A A A G A A T G A T C C T T T A T T A A T G A C A A G G A A A C C N T G N G T
 A A T G C C A C A A T G G C A T A T T G T A A A T G T C A T T T T A A A C A T T G G T A G G C C T T G G T A C A T G A T G C
 T G G A T T A C C T C T C T T A A A A T G A C A C C C T T C C T C G C C T G T T G G T G C T G G C C C T G G G G A G C T N
 G A G C C C A G C A T G C T G G G G A G T G C G G T C T G C T C C A C A C A G T A G T C C C C A N G T G G C C C A N T C C C
 G G C C C A G G C T G C T T T C C G T G T C T T C A G T T C T G T C C A A G C C A T C A G C T C C T T G G G A N T G A T G A
 A C A G A G T C A G A A G C C C A A A G G A A T T G C A N T G T G C A G C A T C A G A N G T A N T N G T C A T A A G T G A
 G A G G C G T G T G T T G A N T G A T T G A C C C A G C G C T T T G G A A A T A A A T G G C A G T G C T T T G T T C A N T T
 A A A G G N C C A A G N T A A A T T T G T A T T G G T T C A T G A T G A A G T C A A A N T G T T A T T C A G A G A T G
 T T T A A T G C A T A T T T A A N T T A T T A A T G T A T T C A T N T C A T G T T T T C T T A T T G T C A C A A G G G T
 A C A G T T A A T G C T G C G T G C T G C T G A A N T C T G T T G G G T G A A N T G G T A T T G C T G

FIGURE 116

GGCCCTTGGGGAGCTGGAGCCCAGCATGCTGGGGAGTGCGGTCAGCTCCACACAGTAGTCCC
CACGTGGCCCACTCCCGGCCAGGCTGCTTTCCGTGTCTTCAGTTCTGTCCAAGCCATCAGC
TCCTTGGGACTGATGAACAGAGTCAGAAGCCCCAAAGGAATTGCACTGTGGCAGCATCAGACG
TACTCGTCATAAGTGAGAGGCGTGTGTTGACTGATTGACCCAGCGCTTTGGAAAATAAATGGC
AGTGCTTTGTTCACTTAAAGGGACCAAGCTAAATTTGTATTGGTTCATGTAGTGAAGTCAAA
CTGTTATTCAGAGATGTTTAAATGCATATTTAACTTATTTAATGTATTTTCATCTCATGTTTTT
TTATTGTCACAAGAGTACAGTTAATGCTGCGTGCTGCTGAACTCTGTTGGGTGAACTGGTAT
TGCTGCTGGAGGGCTGTGGGCTCCTCTGTCTCTGGAGAGTCTGGTCATGTGGAGGTGGG

100
90
80
70
60
50
40
30
20
10
0
100
90
80
70
60
50
40
30
20
10
0

FIGURE 117

CGGAGCTCCGGGTGCTGTGGCCCGGCCTTGGCGGGCGGCCTCCGGCTCAGGCTGGCTGAGA
GGCTCCCAGCTGCAGCGTCCCCGCCGCCTCCTCGGGAGCTCTGATCTCAGCTGACAGTGCC
CTCGGGGACCAAAACAGCGTGGCAGGGTCTCAGTTTGTGCCAGGCTGGAGTTTCACTGCCA
TGATCATGGTTTACTGCAGCCTTGACCTCCTGGGTTCAAGCGATCCTGCTGAGTAGCTGGGA
CTACAGGACAAAATTAGAAATCAAAATGGAAAATATGCTGCTTTGGTTGATATTTTTCACC
CCTGGGTGACCCCTCATTGATGGATCTGAAATGGAATGGGAATTTATGTGGCATTGAGAAA
GGTACCCCGGATTGTCACTGAAAGGACTTTCCATCTCACCAGCCCCGATTGAGGCAGATG
CTAAGATGATGGTAAATACAGTGTGTGGCATCGAATGCCAGAAAGAACTCCCAACTCCCCAGC
CTTTCTGAATTGGAGGATTATCTTTCTATGAGACTGTCTTTGAGAAATGGCACCAGGAACTT
AACCAGGGTGAAAGTTCAAGATTGGTTCTTGAGCCGACTCAAATATCACCACAAAGGGAG
TATCTGTTAGGAGAAAGAGACAGGTGTATGGCACCAGCAGAGTTTCACTCTTGGACAAA
AGGTTCTTAACCAATTTCCCTTTACGACAGCTGTGAAGCTTTCCACGGGCTGTAGTGGCAT
TCTCATTTCCCTCAGCATGTTCTAACTGCTGCCACTGTGTTCATGATGAAAGGACTATG
TCAAAGGGAGTAAAAAGCTAAGGTTAGGGTTGTTGAAGATGAGGAATAAAAGTGGAGGCAAG
AAACGTCGAGGTTCTAAGAGGAGCAGGAGAGAAGCTAGTGGTGGTGACCAAAAGAGAGGGTAC
CAGAGAGCATCTCAGGAGAGAGCGAAGGGTGGGAGAGAAGAAAAAATCTGGCCGGGGT
AGAGGATTGCCAAGGGAGGCCCTTCTTTCACTGGACCCGGGTCAAGAATACCCACATTCCG
AAGGGCTGGGCACGAGGAGGCATGGGGACGCTACCTTGGACTATGACTATGCTCTTCTGGA
GCTGAAGCGTGCTCACAAAAAGAAATACATGGAACCTTGAATCAGCCCAACGATCAAGAAAA
TGCCTGGTGAATGATCCACTTCTCAGGATTGATAACGATAGGGCTGATCAGTTGGTCTAT
CGGTTTTGCAGTGTGTCCGACGAATCCAATGATCTCCTTTACCAATACTGCGATGCTGAGTC
GGGCTCCACCGGTTTCGGGGGTCTATCTGCGCTGGAAGACTCAGACAAAAGAAATGGAAGC
GCAAAATCATTCGGGTCTACTCAGGGCACCAGTGGGTGGATGTCACCGGGGTTCAAGAGGAC
TACAACGTTGCTGTTTCGATCACTCCCTTAAAAACGCCCAGATTGCGCTCTGGATTACGG
GAACGATGCCAATTGTGCTTACGGCTAAACAGAGACCTGAAACAGGGCGGTGATCATCTAAA
TCACAGAGAAAACAGCTCTGCTTACCGTAGTGAGATCACTTCATAGGTTATGCTCGGACTT
GAACTCTGTCAATAGCATTTTCAACATTTTCAAAATCAGGAGATTTTCGTCCATTTAAAAAA
TGTATAGGTGCAGATATTGAAACTAGGTGGGCACCTTCAATGCCAAGTATATACTCTTCTTTA
CATGGTGATGAGTTTCATTTGTAGAAAAATTTTGTGCTTCTTAAAAATTAGACACACTTT
AAACCTTCAACAGGTATTATAAATAACATGTGACTCTTAAATGGACTTATTCTCAGGGTCC
TACTCTAAGAAGAATCTAATAGGATGCTGGTTGTGATTAATATGGAATTCATAGATAAA
GGTAGATGGTAAAGCAATTAGTATCAGAAATAGAGACAGAAAGTTACAAACAGATTGTACTA
CTCTGAGATGGATCCATTCACTCATGCCCTCAATGTTTATATTGTGTTTCTGTTGGGTCT
GGGACATTTAGTTTAGTTTTTTTGAAGAATTACAAATCAGAAAGAAAAGCAAGCATTATAAA
CAAAACTAATACTGTTTTTACTGCTTTAAGAAATAACAATTACAATGTGTATTATTTAAAAA
TGGGAGAAATAGTTTGTCTATGAATAAACCTAGTTTGAAGAAATAGGGAAGCTGAGACATT
TAAGATCTCAAGTTTTTATTTAACTAATACTCAAATATGGACTTTTCATGTATGCATAGGG
AAGACATTCACAAATTATGAATGATCATGTGTTTGAAGCCACATTATTTATGCTATACAT
TCTATGTATGAGGTGCTACATTTTATAGGACAAAGAAATCTGTAATCTTTTCAAGAAAGAGT
CTTTTTCTCCTTGACAAAATCCAGCTTTTGTATGAGGACTATAGGGTGAATTTCTGATTAG
TAATTTTAGATATGCTCTTCTTCTAAAAATGAATAAAATTTATGAATATGA

FIGURE 118

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57253

<subunit 1 of 1, 413 aa, 1 stop

<MW: 47070, pI: 9.92, NX(S/T): 3

MENMLLWLIFFTPGWTLIDGSEMEWDFMWHLRKVPRIVSERTFHLTSPAFEADAKMMVNTVC
GIECQKELPTFSLSELEDYLSYETVPENGTRTLTRVKVQDLVLEPTQNITTKGVSVRKRQV
YGTDSRFSILDKRFLTNPPFSTAVKLSTGCSGILISPQHVLTAACVHDGKDYVKGSKKLRV
GLLKMRNKSGGKKRRGSKRSRREASGGDQREGTREHLQERAKGRRRKKSGRGQRIAEGRPS
FQWTRVKNTHIPKGWARGGMDATLDYDYALLELKRAHKKKYMELGISPTIKKMPGGMIHFS
GFDNDRADQLVYRFCVSDSESNLLYQYCDAESGSTGSGVYLRLKDPDKKNWKRKIIAVYSG
HQWVDVHGVQKDYNAVAVRITPLKYAQICLWIHGNDANCAYG

Important features:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 90-93, 110-113 and 193-196

Glycosaminoglycan attachment site.

amino acids 236-239

Serine proteases, trypsin family, histidine active site.

amino acids 165-170

FIGURE 119

AATGTGAGAGGGGCTGATGGAAGCTGATAGGCAGGACTGGAGTGTAGCACCAGTACTGGAT
GTGACAGCAGGCAGAGGAGCACTTAGCAGCTTATTAGTGTCCGATTCTGATTCCGGCAAGG
ATCCAAGCATGGAATGCTGCCGTCGGGCAACTCCTGGCACACTGCTCCTCTTTCTGGCTTTC
CTGCTCCTGAGTTCAGGACCGCACGCTCCGAGGAGGACCGGACGGCCTATGGGATGCCTG
GGGCCCATGGAGTGAATGCTCACGCACCTGCGGGGAGGGGCTCCTACTCTCTGAGGCGCT
GCCTGAGCAGCAAGAGCTGTGAAGGAAGAAATATCCGATACAGAACATGCAGTAATGTGGAC
TGCCCAACCAGAAGCAGGTGATTTCCGAGCTCAGCAATGCTCAGCTCATAATGATGTCAAGCA
CCATGGCCAGTTTTATGAATGGCTTCCTGTGTCTAATGACCTGACAACCCATGTTCACTCA
AGTGCCAAGCCAAAGGAACAACCCCTGGTTGTTGAACTAGCACCTAAGGCTTTAGATGGTACG
CGTTGCTATACAGAATCTTTGGATATGTGCATCAGTGGTTTATGCCAAATGTTGGCTGCGA
TCACCAGCTGGGAAGCACCCGTCAAGGAAGATAACTGTGGGGTCTGCAACGGAGATGGGTCCA
CCTGCGGGCTGGTCCGAGGGCAGTATAAATCCCAGCTCTCCGCAACCAATCGGATGATACT
GTGGTTGCACTTCCCTATGGAAGTAGACATATTCGCCCTGTCTTAAAGGTCTGTATCACTT
ATATCTGGAACCAAACCCCTCCAGGGGACTAAAGGTGAAAAAGTCTCAGCTCCACAGGAA
CTTTCTCTGTGGACAATTCTAGTGTGGACTTCAGAAATTTCCAGACAAAGAGATACTGAGA
ATGGCTGGACCACTCACAGCAGATTTCAATTGTCAAGATTCGTAACCTCGGGCTCCGCTGACAG
TACAGTCCAGTTCATCTTCTATCAACCCATCATCCACCGATGGAGGGAGACGGATTTCTTTC
CTTGCTCAGCAACCTGTGGAGGAGTTATCAGCTGACATCGGCTGAGTGCTACGATCTGAGG
AGCAACCGTGTGGTTGCTGACCAATACTGTCACTATTACCCAGAGAACATCAAACCCAAACC
CAAGCTTCAGGAGTGCAACTTGGATCCTTGTCCAGCCAGTGACGGATACAAGCAGATCATGC
CTTATGACCTCTACCATCCCTTCTCCTCGTGGGAGGCCACCCCATGGACCGCGTGTCTCTCC
TCGTGTGGGGGGGCATCCAGAGCCGGGCAGTTTCTGTGTGGAGGAGACATCCAGGGGCA
TGTCACTTCAGTGGAAAGAGTGGAAATGCATGTACACCCCTAAGATGCCCATCGCGCAGCCCT
GCAACATTTTTGACTGCCCTAAATGGCTGGCACAGGAGTGTCTCCGTGCACAGTGACATGT
GGCCAGGGCCTCAGATACCGTGTGGTCTCTGCATCGACCATCGAGGAATGCACACAGGAGG
CTGTAGCCCCAAAAACAAGCCCCACATAAAAGAGGAATGCATCGTACCCACTCCCTGCTATA
AACCCAAAGAGAACTTCAGTCGAGGCCAAGTTGCCATGGTTCAAACAAGCTCAAGAGCTA
GAAGAAGGAGCTGCTGTGTGAGGAGGCCCTCGTAAGTTGTAAGAGCAGAGCTGTTCTATA
TTTGAAACTGTTTTGTTTAAAGAAAGCAGTGTCTCACTGGTTGTAGCTTTCATGGGTTCTGA
ACTAAGTGTAAATCATCTACCAAAGCTTTTGGCTCTCAAATTAAGATTGATTAGTTTCAA
AAAAAAAAA

FIGURE 120

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58847

<subunit 1 of 1, 525 aa, 1 stop

<MW: 58416, pI: 6.62, NX(S/T): 1

MECCRRATPGTLLFLAFLLSSRTARSEEDRDGLWDAWGPWSECSRTC GGGASYSLRRLCS
SKSCEGRNIRYRTCSNVDCPPEAGDFRAQQCSAHNDVKHHGQFYEWLPVSNPDNP CSLKCQ
AKGTTLVVELAPKVLDTGTRCYTESLDMCISGLCQIVGCDHQLGSTVKEDNCGVCNGDGSTCR
LVRGQYKSQLSATKSDDTVVALPYGSRHRLVLKGPDLHLYLETKTLQGTKGENSLSSSTGTFL
VDNSSVDFQKFPDKEILRMAGPLTADFIKIRNSGSADSTVQFIFYQPIIHRWRETDFFPCS
ATCGGGYQLTSAECYDLRSNRVVADQYCHYYPENIKPKPKLQECNLDPCPASDGYKQIMPYD
LYHPLPRWEATPWTACSSSCGGGIQSRAVSCVEEDIQGHVTSVEEWKCMYTPKMPIAQPCNI
FDCPKWLAQEWSPCTVTGCGQLRVRVLCIDHRGMHTGGCSPKTKPHIKEECIVPTPCYKPK
EKL PVEAKLPWFKQAQELEEAAVSEEPS

Important features:

Signal peptide:

amino acids 1-25

N-glycosylation site.

amino acids 251-254

Thrombospondin 1

amino acids 385-399

von Willebrand factor type C domain proteins

amino acids 385-399, 445-459 and 42-56

FIGURE 121

CGGACGCGTGGGCGGCGGCTGCGGAACCTCCCGTGGAGGGGCCGCTGGGCCCTCGGGCCTGAC
AGATGCGCAGTGGCCACTGCGGCGGCAGTACTGGCCGCTCTGGGCGGGGCGCTGTGGCTGGCG
GCCCCCGGTTCTGTGGGGCCAGGGTCCAGCGGCTGCGCAGAGGCGGGGACCCCGGCCTCAT
GCACGGGAAGACTGTGCTGATCACCGGGGCGAACAGCGGCCCTGGGCCGCGCCAGCGCCGCC
AGCTACTGCGCCTGGGAGCGCGGGTGATCATGGGCTGCCGGGACCGCGCGCGCCGAGGAG
GCGGCGGGTCAGCTCCGCGCGAGCTCCGCCAGGCCGCGGAGTGGCGCCAGAGCCTGGCGT
CAGCGGGGTGGCGAGCTCATAGTCCGGGAGCTGGACCTCGCCTCGCTGCGCTCGGTGCGCG
CCTTCGCCAGGAAATGCTCCAGGAAGAGCCTAGGCTGGATGTCTTGATCAATAACGCAGGG
ATCTTCCAGTGCCCTTACATGAAGACTGAAGATGGGTTTGAGATGCAGTTCGGAGTGAACCA
TCTGGGGCACTTTCTACTCACCAATCTTCTCCTTGGACTCCTCAAAAGTTGAGCTCCAGCA
GGATTGTGGTAGTTTCTTCCAAACTTTATAAATACGGAGACATCAATTTTGATGACTTGAAC
AGTGAACAAAGCTATAATAAAAGCTTTTGTATAGCCGGAGCAAACCTGGCTAACATTCTTTT
TACCAGGGAAC TAGCCCGCGCTTAGAAGGCACAAATGTCACCGTCAATGTGTGCAATCCTG
GTATTGTACGGACAAATCTGGGGAGGCACATACACATTCCACTGTTGGTCAAACCACTCTTC
AATTTGGTGT CATGGGCTTTTTCAAAACCTCAGTAGAAGGTGCCAGACTTCCATTTATTT
GGCCTCTT CACCTGAGGTAGAAGGAGTGT CAGGAAGATAC TTTGGGGATTGTAAAGAGGAAG
AACTGTTGCCCAAAGCTATGGATGAATCTGTTGCAAGAAAACCTGGGATATCAGTGAAGTG
ATGTTTGGCCTGCTAAAAATAGGAACAAGGAGTAAAAGAGCTGTTTATAAACTGCAATATCAG
TTATATCTGTGATCAGGAATGGTGTGGATTGAGAACTTGTTACTTGAAGAAAAAGAATTTTG
ATATTGGAATAGCCTGCTAAGAGGTACATGTGGGTATTTTGGAGTTACTGAAAAATTTATTTT
TGGGATAAGAGAAATTT CAGCAAAGATGTTTTAAATATATATAGTAAGTATAATGAATAATAA
GTACAATGAAAAATACAATTATATTGTAAAATTATAACTGGGCAAGCATGGATGACATATTA
ATATTGTGAGAATTAAAGTACTCAAAGTGCTATCGAGAGGTTTTTCAAGTATCTTTGAGTT
TCATGGCCAAAGTGTTAACTAGTTTTACTACAATGTTTGGTGTGTGTGTGGAATTATCTGC
CTGGTGTGTGCACACAAGTCTTACTTGGAATAAAATTACTGGTAC

FIGURE 122

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58747

<subunit 1 of 1, 336 aa, 1 stop

<MW: 36865, pI: 9.15, NX(S/T): 2

MAVATAAAVLAALGGALWLAARRFVGPRVQRLRRGGDPGLMHGKTVLITGANSGLGRATAAE
LLRLGARVIMGCRDRARAEEAAGQLRREL RQAAECGPEPGVSGV GELIVRELDLASLRSVRA
FCQEMLQE EPRLDVLINNAGIFQCPYMKTEDGFEMQPGVNHLGHFLLTNLLLGLLKSSAPSR
IVVVSSKLYKYGDINFDDLNSEQSYNKSPCYSRSKLANILFTRELARRLEGTNVTNVNLHPG
IVRTNLGRHIHIPLLVKPLFNLVSWAFFKTPVEGAQTSIYLASSPEVEGVSGRYFGDCKEEE
LLPKAMDES VARKLWDISEVMVGLLK

Important features:

Signal peptide:

amino acids 1-21

Short-chain alcohol dehydrogenase family protein

amino acids 134-144, 44-56 and 239-248

N-glycosylation site.

amino acids 212-215 and 239-242

FIGURE 123

GGGGATTGTAAAGAGGAAGNACTGTGCCCAAAGNTATGGATGAATCTGTTGCAAGAAAATTN
TGGGATATCAGTGAAGTGATGGTTNGCCTGCTAAAAATAGGAACAAGGAGTAAAAGAGCTGTT
TATAAACTGCATATCAGTTATATCTGTGATCAGGAATGGTGTGGATTGAGAACTTGTTACT
TGAAGAAAAAGAATTTTGATATTGGAATAGCCTGNTAAGAGGNACATGTGGGTATTTTGGAG
TTACTGAAAAATTATTTTGGGATAAGAGAATTCAGCAAAGATGTTTTAAATATATATAGT
AAGTATAATGAATAATAAGTACAATGAAAAATACAATTATATTGTAAAAATTATACTGGGCA
AGCATGGATGACATATTAATATTTGTGAGAATTAAGTGACTIONAAAGTGCATCGAGAGGTTT
TTCAAGTATCTTTGAGTTTCATGGCCAAAGTGTTAACTAGTTTTACTACAATGTTTGGTGTT
TGTTGTGAAATTATCTGCCTGGCTT

FIGURE 124

GAGAGGACGAGGTGCCGCTGCCTGGAGAATCCTCCGCTGCCGTGCGCTCCCGAGCCAGCC
CTTTCCTAACCCAAACCAACCTAGCCAGTCCCAGCCGCGAGCGCTGTCCCTGTACAGGAC
CCCAGCGTTACCATGCGATCCTGCCGTCTTCTATCCTTACCCGACCTCAGATGCTCCCTTCT
GCTCCTGGTAACTTGGGTTTTTACTCTGTAACTGAAATAACAAGTCTTGCTACAGAGA
ATATAGATGAAATTTTAAACAATGCTGATGTTGCTTTAGTAAATTTTATGCTGACTGGTGT
CGTTTCAGTCAGATGTTGCATCCAATTTTGAGGAAGCTTCCGATGTCATTAAAGGAAGAATT
TCCAATGAAATCAAGTAGTGTGTCAGAGTTGATTGTGATCAGCACTCTGACATAGCCC
AGAGATACAGGATAAGCAAATACCAACCTCAAATGTTTCGTAATGGGATGATGATGAAG
AGAGAATACAGGGGTCAGCGATCAGTGAAGCATTGGCAGATTACATCAGGCAACAAAAAG
TGACCCCATTCAGAAATTCGGGACTTAGCAGAAATCACCCTCTGATCGCAGCAAAAGAA
ATATCATTGGATATTTTGAAGAAAGGACTCGGACAACATAGAGTTTTTGAACGAGTAGCG
AATATTTTGCATGATGACTGTGCCTTCTTTCTGCATTGGGGATGTTTCAAAACCGGAAAG
ATATAGTGGCGACAACATAATCTACAAACCACCGGGCATTCTGCTCCGGATATGGTGTACT
TGGGAGCTATGACAAATTTGATGTGACTTACAATTGGATTCAAGATAAATGTTCTCTCTT
GTCCGAGAAATAACATTTGAAATGGAGAGGAATTGACAGAAGAAGGACTGCCTTTTCTCAT
ACTCTTTCATGAAAGAAGATACAGAAAGTTTAGAAATATTCGAAATGAAGTAGCTCGGC
AATTAATAAGTGAAAAAGGTACAATAAACTTTTACATGCCGATTGTGACAAATTTAGACAT
CCTCTTCTGCACATACAGAAAACCTCCAGCAGATTGCTGTAAATCGCTATTGACAGCTTTAG
GCATATGTATGTGTTTGGAGACTTCAAAGATGTATTAATTCCTGGAAAACCTCAAGCAATTG
TATTTGACTTACATCTCGGAAAACCTGCACAGAGAATTCATCATGACCTGACCCAACTGAT
ACAGCCCCAGGAGAGCAAGCCCAAGATGTAGCAAGCAGTCCACCTGAGAGCTCCTTCCAGAA
ACTAGCACCCAGTGAATATAGGTATACTCTATTGAGGGATCGAGATGAGCTTTAAAAAACTTG
AAAAACAGTTTGTAAAGCCTTTCAACAGCAGCATCAACCTACGTGGTGGAAATAGTAAACCTA
TATTTTCATAATTCTATGTGTATTTTTATTTTGAATAACAGAAAGAAATTTAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 125

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57689
<subunit 1 of 1, 406 aa, 1 stop
<MW: 46927, pI: 5.21, NX(S/T): 0
MHPAVFLSLPDLRCSLLLLVTWVTFPVTTEITSLATENIDEILNNADVALVNFYADWCRFSQ
MLHPIFEEASDVIKEEFPNENQVVFARVDCDQHSDIAQRYRISKYPTLKLFRNGMMMKREYR
GQRSVKALADYIRQQKSDPIQEIRDLAEITTLDRSKRNIIGYFEQKSDNRYRVFERVANILH
DDCAFLSAFGDVSKPERYSGDNIIYKPPGHSAPDMVYLGAMTNFDTVYTNWIQDKCVPLVREI
TFENGEELTEEGLPFLILFHMKEDESLEIFQNEVARQLISEKGTINFLHADCDKFRHPLLH
IQKTPADCPVIAIDSFRHMYVFGDFKDVLI PGKLGQFVFDLHSGKLHREFHHGPDPTDTPAG
EQAQDVASSPPESSFQKLAPSEYRYTLLRDRDEL

Important features:

Signal peptide:

amino acids 1-29

Endoplasmic reticulum targeting sequence.

amino acids 403-406

Tyrosine kinase phosphorylation site.

amino acids 203-211

Thioredoxin family proteins

amino acids 50-66

FIGURE 127

AGAGGCCTCTCTGGAAGTTGTCCCGGGTGTTCGCCGNGGAGCCCGGGTCGAGAGGACNAGG
TGCCGCTGCCTGGAGAATCCTCCGCTGCCGTCGGCTCCCGAGCCCAGCCCTTTCCTAACCC
AACCCAACCTAGCCCNGTCCCAGCCGCCAGCGCCTGTCCCTGTCNCGGANCCCAGCGTNACC
ATGCATCCTGCCGTCTTCCTATCCTTACCCGACCTCAGATGCTCCCTTCTGCTCCTGGTAAC
TTGGGTTTTTACTCCTGTAACAACCTGAAATAACNNGTCTTGATACNNAGAATATAGATGAAA
TTTTAAACNATGCTGATGTGGCTTTAGTCAATTTTATGCTGACTGGTGTGTTTTAGTCAG
ATGTGGCATCCAATTTTGGAGGANGCTTCGGATGTCATTAAGGAAGAATTTCCAAATGAAAA
TCAAGTAGTGTGTTGCCAGAGTTGATTGTGATCAGCACTCTGACATAGCCCAGAGATACAGGA
TAAGCAAATACCCAACCTCAAATTGTTTCGTAATGGGATGATGATGAAGAGAGAATACAGG
GGTCAGCGATCAGTGAAAGCATTGGCAGATTACATCAGGC

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 128

GCCCACGCGTCCGATGCGGTTACAGTTCGCGGCTTCTGCTACATGCTGGCGCTGCTGCTCA
CTGCCGCGCTCATCTTCTTCGCCATTTGGCACATTATAGCATTGTAGAGCTGAAGACTGAT
TACAAGAATCCTATAGACCAGTGTAATACCCTGAATCCCCCTGTACTCCCAGAGTACCTCAT
CCACGCTTTCTTCTGTGCATGTTTCTTTGTGCAGCAGAGTGGCTTAACTGGGTCTCAATA
TGCCCTCTTGGCATATCATATTTGGAGGTATATGAGTAGACCAGTGATGAGTGGCCAGGA
CTCTATGACCTACAACCATCATGAATGCAGATATTCTAGCATATTGTCAGAAGGAAGGATG
GTGCAAAATTAGCTTTTTATCTTCTAGCATTTTTTTACTACCTATATGGCATGATCTATGTTT
TGGTGAGCTCTTAGAACAACACACAGAAGAATTGGTCCAGTTAAGTGCATGCAAAAAGCCAC
CAAAATGAAGGATTCTATCCAGCAAGATCCTGTCCAAGAGTAGCCTGTGGAATCTGATCAGT
TACTTTAAAAAATGACTCCTTATTTTTTAAATGTTTCCACATTTTTGCTTGTGGAAAGACTG
TTTTCATATGTTATACTCAGATAAAGATTTTAAATGGTATTACGTATAAAATTAATATAAAAT
GATTACCTCTGGTGTGTGACAGGTTTGAACCTTGCACTTCTTAAGGAACAGCCATAATCCTCTG
AATGATGCATTAATTACTGACTGTCTAGTACATTGGAAGCTTTTGTTTATAGGAACCTTGTA
GGGCTCATTTGGTTTCATTGAAACAGTATCTAATTATAAATTAGCTGTAGATATCAGGTGC
TTCTGATGAAGTGAAATGTATATCTGACTAGTGGGAACTTCATGGGTTTCCTCATCTGTC
ATGTCGATGATTATATATGGATACATTTACAAAAATAAAAGCGGGAATTTCCCTTCGCTT
GAATATTATCCCTGTATATTGCATGAATGAGAGATTTCCCATATTTCCATCAGAGTAATAAA
TATACTTGCTTTAATTCTTAAGCATAAGTAAACATGATATAAAAAATATATGCTGAATTACTT
GTGAAGAATGCATTTAAAGCTATTTTAAATGTGTTTTATTGTGAAGACATTACTTATTAAG
AAATTGGTTATTATGCTTACTGTTCTAATCTGGTGGTAAAGGTATTCTTAAGAATTTGCAGG
TACTACAGATTTTCAAAACTGAATGAGAGAAAATTGTATAACCATCCTGCTGTTCTTTAGT
GCAATACAATAAACTCTGAAATTAAGACTC

FIGURE 129

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA23330

<subunit 1 of 1, 144 aa, 1 stop

<MW: 16699, pI: 5.60, NX(S/T): 0

MAFTFAAFCYMLALLTAALIFFAIWHIIAFDELKTDYKNPIDQCNTLNPLVLPEYLIHAFF
CVMFLCAAEWLTLGLNMPELLAYHIWRYMSRPVMSGPGLYDPTTIMNADILAYCQKEGWCKLA
FYLLAFFYYLYGMIYVLVSS

Important features:

Signal peptide:

amino acids 1-20

Type II transmembrane domain:

amino acids 11-31

Other transmembrane domain:

amino acids 57-77 and 123-143

FIGURE 130

ATTATAGCATTTGATGAGCTGAAGACTGATTACAAGATCCTATAGACCAGTGTAATACCCTG
AATCCCCCTTGACTCCCAGAGTACCTCATCCACGCTTTCCTCTGTGTCATGTTTCTTTGTGC
AGCAGAGTGGCTTACACTGGGTCTCAATATGCCCTCTTGGCATATCATATTTGGAGGTATA
TGAGTAGACCAGTGATGAGTGGCCCAGGACTCTATGACCCTACAACCATCATGAATGCAGAT
ATTCTAGCATATTGTCAGAAGGAAGGATGGTGCAAATTAGCTTTTTATCTTCTAGCATTTTT
TTACTACCTATATGGCATGATCTATGTTTTGGTGAGCTCTTAGAACAACACAGAAGAATT
GGTCCAGTTAAGTGCATGCAAAAAGCCACCAAATGAAGGGATTCTATCCAGCAAGATCCTGT
CCAAGAGTAGCCTGTGGAATCTGATCAGTTACTTTAAAAAATG

100
90
80
70
60
50
40
30
20
10
0
100
90
80
70
60
50
40
30
20
10
0

FIGURE 131

CGGACGCGTGGGGAAACCCCTTCGAGAAAAACAGCAACAAGCTGAGCTGCTGTGACAGAGGG
GAACAAGATGGCGGCGCCGAAGGGGAGCCTCTGGGTGAGGACCCAACTGGGGCTCCCGCCGC
TGCTGCTGCTGACCATGGCCTTGGCCGGAGGTTCTGGGACCGCTTCGGCTGAAGCATTTGAC
TCGGTCTTGGGTGATACGGCGTCTTGCCACCGGGCCTGTGAGTTGACCTACCCCTTGCACAC
CTACCCTAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTGCAGGCTGTTTTCAATTTGTC
AGTTTGTGGATGATGGAATTGACTTAAATCGAACTAAATTGGAATGTGAATCTGCATGTACA
GAAGCATATTTCCCAATCTGATGAGCAATATGCTTGCCATCTTGGTTGCCAGAATCAGCTGCC
ATTCGCTGAACTGAGACAAGAACAACCTTATGTCCCTGATGCCAAAAATGCACCTACTCTTTC
CTCTAACTCTGGTGAGGTCATTCTGGAGTGACATGATGGACTCCGCACAGAGCTTCATAACC
TCTTCATGGACTTTTTATCTTCAAGCCGATGACGAAAAATAGTTATATTTCCAGTCTAAGCC
AGAAATCCAGTACGCACCCACATTTGGAGCAGGAGCCTACAAATTTGAGAGAATCATCTCTAA
GCAAAATGTCCATCTCTGCAAATGAGAAATTCACAAGCGCACAGGAATTTTCTTGAAGATGGA
GAAAGTGATGGCTTTTTAAGATGCCTCTCTCTTAACTCTGGGTGGATTTTAACTACAACTCT
TGTCTCTCGGTGATGGTATTGCTTTGGATTGTTGTGTGCAACTGTTGTACAGCTGTGGAGC
AGTATGTTCCCTCTGAGAAGCTGAGTATCTATGGTGACTTGGAGTTTATGAATGAACAAAAG
CTAAACAGATATCCAGCTTCTTCTTGTGGTTGTTAGATCTAAAACTGAAGATCATGAAGA
AGCAGGGCCTCTACCTACAAAAGTGAATCTTGCTCATTTCTGAAATTTAAGCATTTTCTTTT
AAAAGACAAGTGTAATAGACATCTAAAATTCACCTCCTCATAGAGCTTTTAAAATGGTTTCA
TTGGATATAGGCCTTAAGAAATCACTATAAAATGCAAATAAAGTTACTCAAATCTGTG

FIGURE 132

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA26847

<subunit 1 of 1, 323 aa, 1 stop

<MW: 36223, pI: 5.06, NX(S/T): 1

MAAPKGS LWVRTQLGLPPLLLLTMALAGGSGTASAEAFDSVLGDTASCHRA CQLTYPLHTYP
KEEELYACQRCRLFSICQFVDDGIDLNR TKLECESACTEAYSQSDEQYACHLGCQNQLPFA
ELRQEQ LMSLMPKMHLLFPLTLVRSFWS DMMSAQSFITSSWTFYLQADDGKIVIFQSKPEI
QYAPHLEQEPTNLRESSLSKMSY LQMRNSQAHRNFLEDGESDGLRCLSLNSGWILTTTLVL
SVMVLLWICCATVATAVEQYVPSEKLSIYGDLEFMNEQKLNRYPASSLVVVRSKTEDHEEAG
PLPTKVNLAHSEI

Important features:

Signal peptide:

amino acids 1-31

Transmembrane domain:

amino acids 241-260

N-glycosylation site.

amino acids 90-93

[illegible]

FIGURE 134

Journal of Management Education 30(6)p.789-804
© The Author(s) 2006
Reprints and permissions:
<http://www.sagepub.com/journalsPermissions.nav>

FIGURE 135

GCGAGGTGGCGATCGCTGAGAGGCAGGAGGGCCGAGGCGGGCCTGGGAGGCGGCCCGGAGGT
GGGGCGCCGTGGGGCCGGCCCGCACGGGCTTCATCTGAGGGCGCACGGCCCCGACCGAGC
GTGCGGACTGGCCTCCCAAGCGTGGGGCGACAAAGCTGCCGAGACTGCAATGGGGCCGCGGTG
GGGATTCTTGTGCTTGGCCTCCTGGGCGCCGTGTGGCTGCTCAGCTCGGGCCACGGAGAGGAGC
AGCCCCCGGAGACAGCGGCACAGAGGTGCTTCTGCCAGGTTAGTGTTACTTGGATGATTGT
ACCTGTGATGTTGAAACCATTGATAGATTTAATAACTACAGGCTTTTCCCAAGACTACAAAA
ACTTCTTGAAAGTGACTACTTTAGGTATTACAAGGTAAACCTGAAGAGGCCGTGCTCTTCT
GGAATGACATCAGCCAGTGTGGAAGAAGGGACTGTGCTGTCAAACCATGTCAATCTGATGAA
GTTCTGATGGAATTAATCTGCGAGCTACAAGTATTCTGAAGAAGCCAATAATCTCATTGA
AGAATGTGAACAAGCTGAACGACTTGGAGCAGTGGATGAATCTCTGAGTGAGGAAACACAGA
AGGCTGTTCTTCAGTGGACCAAGCATGATGATTCTTCAGATAACTTCTGTGAAGCTGATGAC
ATTCACTCCCTGAAGCTGAATATGTAGATTTGCTTCTTAATCCTGAGCGCTACACTGGTTA
CAAGGGACAGATGCTTGGAAAAATATGGAATGTCATCTACGAAGAAAACCTGTTTTAAGCCAC
AGACAATTAAGACCTTTAAATCCTTTGGCTTCTGGTCAAGGGACAAGTGAAGAGAACACT
TTTTACAGTTGGCTAGAAGGTCTCTGTGTAGAAAAAGAGCATTCTACAGACTTATATCTGG
CCTACATGCAAGCATTAATGTGCATTTGAGTGCAGATATCTTTACAAGAGACCTGGTTAG
AAAAGAAATGGGGACACAACATTACAGAATTTCAACAGCGATTTGATGGAATTTTGACTGAA
GGAGAAGGTCCAAGAAGGCTTAAGAACTTGATTTTTCTCTACTTAATAGAACTAAGGGCTTT
ATCCAAAGTGTTACCATTTCTTCGAGCGCCAGATTTTCAACTCTTTACTGAAATAAAATTC
AGGATGAGGAAAAACAAATGTTACTTCTGGAAATACTTCATGAAATCAAGTCATTTTCTTTG
CATTTTGATGAGAATTCATTTTTGCTGGGGATAAAAAAGAGCACACAACTAAAGGAGGA
CTTTCGACTGCATTTTAGAAAATTTCAAGAATTATGGATTGTGTTGGTTGTTTTAAATGTC
GTCTGTGGGAAAGCTTCAGACTCAGGGTTTGGGCACTGCTCTGAAGATCTTATTTTCTGAG
AAATTGATAGCAATATGCCAGAAAGTGGACCTAGTTATGAATTCATCTAACCAGACAAGA
AATAGTATCATTATTCAACGATTTGGAAGAATTTCTACAAGTGTGAAGAATTAGAAAACT
TCAGGAACCTGTTACAGAATATTCATTAAAGAAAAACAGCTGATATGTGCTGTTTCTGGAC
AATGGAGGCGAAAGAGTGAATTTCAATCAAAGGCATAATAGCAATGACAGTCTTAAGCCAA
ACATTTTATATAAAGTTGCTTTTGTAAAGGAGAATTATATGTTTTAAGTAAACACATTTTT
AAAAATTGTGTTAAGTCTATGTATAATACTACTGTGAGTAAAGTAATACTTTAATAATGTG
GTACAAATTTTAAAGTTAATATTGAATAAAAGGAGGATTATCAAATTAATAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 136

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53974

<subunit 1 of 1, 468 aa, 1 stop

<MW: 54393, pI: 5.63, NX(S/T): 2

MGRGWGFLFGLLGAVWLLSSGHGEEQPPETAQAQRFCQVSGYLDCTCDVETIDRFNNYRLF
PRLQKLLESDFRYYKVNLRKPCPFWNDISQCGRRDCAVKPCQSDEVDPGDKSASYKYSEEA
NNLIEECEQAERLGAVDESLSEETQKAVLQWTKHDDSSDNFCEADDIQSPEAEYVDLLLNPE
RYTGYKGPDAWKIWNVIYEENCFKPQTIKRPLNPLASGQGTSEENTFYSWLEGLCVEKRAFY
RLISGLHASINVHLSARYLLQETWLEKKWGHNITEFQQRFDGILTEGEGPRRLKNLYFLYLI
ELRALSKVLPPFFERPDLFTGNKIQDEENKMLLLEILHEIKSFPLHFDENSFFAGDKKEAH
KLKEDFRLHFRNISRIMDCVGCFCRLWGKLQTQGLGTALKILFSEKLIANMPESGSPSYEFH
LTRQEIVSLFNAFGRISTSVKELENFRNLLQNIH

Important features:

Signal peptide:

amino acids 1-23

N-glycosylation site.

amino acids 280-283 and 384-387

Amidation site.

amino acids 94-97

Glycosaminoglycan attachment site.

amino acids 20-23 and 223-226

Aminotransferases class-V pyridoxal-phosphate

amino acids 216-222

Interleukin-7 proteins

amino acids 338-343

FIGURE 137

GCTGGAAATATGGATGTCATCTACGAGAACTGTTTTAAGCCACAGACAATTAAAAGACCTT
TAAATCCTTTGGCTTCTGGTCAAGGGACAAGTGAAGAGNACACTTTTTACAGTTGGCTAGAA
GGTCTCTGTGTAGAAAAAAGAGCATTCTACAGACTTATATCTGGCCTACATGCAAGCATTAA
TGTGCATTTGAGTGCAAGATATCTTTTACAAGAGACCTGGTTAGAAAAGAAATGGGGACACA
ACATTACAGAATTTNAACAGCGATTTGATGGAATTTTGACTGAAGGAGAAGGTCCAAGAAGG
CTTAAGAACTTGTATTTTCTCTACTTAATAGAACTAAGGGCTTTATCCAAAGTGTACCATT
CTTNGAGCGCCAGATTTTCAACTNNTTACTGGAAATAAAATTCAGGATGAGGNAAACAAA
TGTTACTTTTGGAAATACTTCATGAAATCAAGTCATTTCCTTTGCATTTTGTATGAGAATTCA
TTTTTTTGCTG

100
99
98
97
96
95
94
93
92
91
90
89
88
87
86
85
84
83
82
81
80
79
78
77
76
75
74
73
72
71
70
69
68
67
66
65
64
63
62
61
60
59
58
57
56
55
54
53
52
51
50
49
48
47
46
45
44
43
42
41
40
39
38
37
36
35
34
33
32
31
30
29
28
27
26
25
24
23
22
21
20
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1

FIGURE 138

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGTGGGAGGGGCGAGGATGGGAGGGAA
AGTGAAGAAAAACAGAAAAGGAGAGGGACAGAGGCCAGAGGACTTCTCATACTGGACAGAAAC
CGATCAGGCATGGAACTCCCCCTTCGTCACTCACTGTCTCTTGCCCTGGTGTTCTCTGACAGG
TCTCTGCTCCCCCTTTAACCTGGATGAACATCACCCACGCCTATTCCCAGGGCCACCAGAAG
CTGAATTTGGATACAGTGTCTTACAACATGTTGGGGGTGGACAGCGATGGATGCTGGTGGGC
GCCCCCTGGGATGGGCCTTCAGGCGACCGGAGGGGGGACGTTTATCGCTGCCCTGTAGGGGG
GGCCCAATGCCCCATGTGCCAAGGGCCACTTAGGTGACTACCAACTGGGAAATTCATCTC
ATCCTGCTGTGAATATGCACCTGGGGATGTCTCTGTTAGAGACAGATGGTGATGGGGGATTC
ATGGTGAGCTAAGGAGAGGGTGGTGGCAGTGTCTCTGAAGGTCCATAAAAGAAAAAGAGAA
GTGTGTTAAGGAAAAATGGTCTGTGTGGAGGGGTCAAGGAGTTAAAAACCTAGAAAGCAAA
AGGTAGGTAAATGTGAGGAGTAGTCTTCATGCCTCCTTCAACTGGGAGCATGTTCTGAGGGT
GCCCCCACAAGCCTGGGAGTAACATATTTCCCCCATCCCAGGCCTGTGCCCTCTCTGGTCT
CGTGCTTGTGGCAGCTCTGTCTTCAGTTCTGGGATATGTGCCCGTGTGGATGCTTCATTCCA
GCCTCAGGAAGCCTGGCACCCTGCCCCAACGTGAGCCAGAGGAAGGCTGAGTACTTGGTT
CCCAGAAGGAGATACTGGGTGGGAAAAAGATGGGGCAAAAGCGGTATGATGCCTGGCAAAGGG
CCTGCATGGCTATCCTCATTGCTACCTAATGTGCTTGCAAAGCTCCATGTTTCTTAA CAGA
TTCAGACTCCTGGCCAGGTGTGGTGGCCACACCTGTAATTCTAGCACTTGGGAGGCCAAG
GTGGGCAGATCACTTGAGGTGAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAAACTCCAT
CTCTACTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA
ATCTACTCGGAGGCTAAGACAGGAGACTCTCACTTCAACCCAGGAGGTGGAGGTTGCGGTG
AGCCAAGATTGTGCCTCTGCACTCTAGCGTGGGTGACAGAGTAAGCGAGACTCCATCTCAAA
AATAATAATAATAATAATTCAGACTCCTTATCAGGAGTCCATGATCTGGCCTGGCACAGTAA
CTCATGCCTGTAATCCCAACATTTTGGGAGGCCAACGAGGAGGATTGCTTGAGGTCTGGAG
GTTTGAGACCAGCCTGGGCAACATAGAAAGACCCCATCTCTAAATAAATGTTTTAAAAAT

FIGURE 139

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57039
><subunit 1 of 1, 124 aa, 1 stop
><MW: 13352, pI: 5.99, NX(S/T): 1
MELPFVTHLFLPLVFLTGLCSPPFNLDHHPRLPFGPPEAEFGYSVLQHVGGGQRWMLVGAPW
DGPSGDRRGDVYRCVPVGAHNAPCAKGHLGDYQLGNSSHPAVNMHLGMSLLETGDGGGFMVS

Important features:

Signal peptide:

amino acids 1-22

Cell attachment sequence.

amino acids 70-73

N-glycosylation site.

amino acids 98-101

Integrins alpha chain proteins

amino acids 67-81

$$\begin{aligned} \frac{d}{dt} \left(\frac{1}{2} \dot{\theta}^2 \right) &= \dot{\theta} \ddot{\theta} \\ \frac{d}{dt} \left(\frac{1}{2} \dot{\phi}^2 \right) &= \dot{\phi} \ddot{\phi} \\ \frac{d}{dt} \left(\frac{1}{2} \dot{\psi}^2 \right) &= \dot{\psi} \ddot{\psi} \end{aligned}$$

CACAGTTCACCCACCATCACTCNTCCCATTCCTTCCAACCTTATTTTTAGCTTGCCATTGGGA
GGGGGCAGGATGGGAGGGAAAGTGAAGAAAAAGAGAGGGACAGAGGCCAGAGGAC
TTCTCATACTGGACAGAAACCGATCAGGCATGGAACTCCCCCTTCGTCACTACCTGTTCTTG
CCCCTGGTGTTCTCTGACAGGTCCTGCTCCCCCTTAACTGGATGAACATCACCCACGCCT
ATTCACAGGGCCACAGAAGCTGAATTGGATACAGTGTCTTACAACATGTTGGGGGTGGAC
AGCGATGGATGCTGGTGGGCGCCCCCTGGGATGGGCCTTCAGCGACCGAGGGGGGACGTT
TATCGCTGCCCTGTAGGGGGGGCCCAATGCCCATGTGCCAAGGGCCACTTAGGTGACTA
CCAAGTGGGAAATTCATCTCATCTGCTGTGAATATGCACCTGGGGATGCTCTGTTAGAGA
CAGATGGTGTAGTG

FIGURE 141

AAAGTTACATTTTCTCTGGAACCTCTCCTAGGCCACTCCCTGCTGATGCAACATCTGGGTTTG
GGCAGAAAGGAGGGTGCTTCGGAGCCCGCCCTTTCTGAGCTTCCTGGGCCGGCTCTAGAACA
ATTCAGGCTTCGTCGCACTCAGACCTCAGCTCCAACATATGCATTCTGAAGAAAGATGGCT
GAGATGGACAGAAATGCTTTATTTTGGAAAGAAACAATGTTCTAGGTCAAACAGTGTCTACCA
AATGCAGACTTTACAAATGGTTCTAGAAGAAATCTGGACAAGTCTTTTCATGTGGTTTTTCT
ACGCATTGATTCCATGTTTGTCTACAGATGAAGTGGCCATTCTGCCCTGCCCTCAGAACCTC
TCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCAGTGATCGCGCTCGGAGA
AACAGTGTACTATTCTGTGGAATACCAGGGGGAGTACGAGAGCCGTGTACACGAGCCACATCT
GGATCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCTGTAGTGTGATGTCTGATGACATC
ACGGCCACTGTGCCATACAACCTTCGTGTGAGGGCCACATTGGGCTCACAGACCTCAGCCTG
GAGCATCTGAAGCATCCCTTTAATAGAAACTCAACCATCCTTACCCGACCTGGGATGGAGA
TCACCAAAGATGGCTTCACCTGGTTATTGAGCTGGAGGACCTGGGGCCCCAGTTTGTAGTTT
CTTGTCGCTACTGGAGGAGGGAGCTGGTGCCGAGGAACATGTCAAATGTTGAGGAGTGG
GGGTATTCCAGTGCACTAGAAACCATGGAGCCAGGGCTGCATACTGTGTGAAGGCCAGA
CATTCTGTGAAGGCCATTGGGAGGTACAGCGCCTTCAGCCAGACAGAATGTGTGGAGGTGCAA
GGAGAGGCCATTCCCCCTGGTACTGGCCCTGTTTGCCTTTGTTGGCTTCATGCTGATCCTTGT
GGTCGTGCCACTGTTCTGTCTGAAAAATGGGCCGGCTGCCTCAGTACTCCTGTTGCCCGTGG
TGGTCCTCCAGACACCTTGAATAAACCAATTCACCCAGAAAGTTAATCAGCTGCAGAAGG
GAGGAGGTGGATGCCTGTGCCACGGCTGTGATGTCTCCTGAGGAACTCCTCAGGGCCTGGAT
CTCATAGGTTTTGCGGAAGGGCCAGGTGAAGCCGAGAACCCTGGTCTGCATGACATGGAAACC
ATGAGGGGACAAGTTGTGTTTTCTGTTTTCCGCCACGGACAAGGGATGAGAGAAGTAGGAAGA
GCCTGTTGTCTACAAGTCTAGAAGCAACCATCAGAGGCAGGGTGGTTTTGTCTAACAGAACAC
TGACTGAGGCTTAGGGGATGTGACCTCTAGACTGGGGGCTGCCACTTGTCTGGCTGAGCAACC
CTGGGAAAAGTGACTTCATCCCTTCGGTCTTAAGTTTTCTCATCTGTAATGGGGGAATTACC
TACACACCTGCTAAACACACACACACAGAGTCTCTCTATATATACACACGTACACATAAA
TACACCCAGCACCTGCAAGGCTAGAGGGAACCTGGTGACACTCTACAGCTGACTGATTTCAG
TGTTTCTGGAGAGCAGGACATAAATGTATGATGAGAATGATCAAGGACTCTACACACTGGGT
GGCTTGGAGAGCCCACTTTCAGAGAATAATCCTTGAGAGAAAAGGAATCATGGGAGCAATGG
TGTTGAGTTCACTTCAAGCCCAATGCCGTTGCAGAGGGGAATGGCTTAGCGAGCTCTACAGT
AGGTGACCTGGAGGAAGGTACAGCCACACTGAAAAATGGGATGTGCATGAACACGGAGGATC
CATGAACACTGTAAAGTGTGTACAGTGTGTGCACACTGCAGACAGCAGGTGAAATGTATGT
GTGCAATGCGACGAGAATGCAGAAGTCAGTAACATGTGCATGTTTGTGTGTCTCTTTTCT
TGTTGGTAAAGTACAGAATTGAGCAATAAAAAGGCCACCTTGCCAAAAGCGGTAAAAAA
AAAAA

FIGURE 142

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57033
<subunit 1 of 1, 311 aa, 1 stop
<MW: 35076, pI: 5.04, NX(S/T): 2
MQTFMTMLEEIIWTSLFMWWFFYALIPCLLTDEVAIIPAPQNLSVLSTNMKHLMLWSPVIAPGE
TVYYSVEYQGEYESLYTSHIWIPSSWCSLTEGPECDDITATVPYNLRVRATLGSQTS
SILKHPFNRNSTILTRPGMEITKDGPHLVIELEDLGPQFEFLVAYWRREPGABEHVKMVRSG
GIPVHLETMEPGAAYCVKAQTFVKAIGRYSAFSQTECEVEVQGEAIPVLALFAFVGFMILILV
VVPLFVWKMGRLQLQYSCCPVVVLPDTLKITNSPQKLISCRREEVDACATAVMSPEELLRAWIS

Important features:

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 230-255

N-glycosylation site.

amino acids 40-43 and 134-137

Tissue factor proteins.

amino acids 92-119

Integrins alpha chain proteins

amino acids 232-262

FIGURE 143

TCCTGCTGATGCACATCTGGGTTTGGCAAAAGGAGGTTGCTTCGAGCCGCCCTTTCTAGCTT
CCTGGCCGGCTCTAGAACAATTGAGGCTTCGCTGCGACTAGACCTCAGCTCCAACATATGCA
TTCTGAAGAAAGATGGCTGAGATGACAGAATGCTTTATTTTGAAAGAAACAATGTTCTAGG
TCAAACCTGAGTCTACCAAATGCAGACTTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCT
TTTCATGTGGTTTTTCTACGCATTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGC
CTGCCCTCAGAACCTCTCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCA
GTGATCGCGCCTGGAGAAAACAGTGTACTATTCTGTGGAATACCAGGGGGAGTACGAGAGCCT
GTACACGAGCCACATCTGGATCCCAGCAGCTGGTGTCACTCACTGAAGGTCTGAGTGTG
ATGTCACTGATGACATCACGGCCACTGTGCCATACAACCTTTGTGTGAGGCCACATTGGGC
TCACAGACCTCAGCCTGGAGCATCCTGAAGCATCCCTTTAATAGAACTCAACCATCCTTAC
CCGACCTGGGATGGAGATCACCAAAGATGGCTTNCACCTGGTTATTGAGCTGGAGGACCTGG
GGCCCCAGTTTGAGTTCCTTGTGGCCTANTGGAGGAGGGGCGAACCCCTTGCGGCGCAAGGG
GTNCGGAACCCCTTGCGGCCGCTGGGGTATCTCTCGAGAAAAGAGAGGCCCAATATGACCC
ACATACTCAATATGGACGAANTGCTATTGTCCACCTGTTTGAGTGGCGCTGGGTTGAT

FIGURE 144

CCCACGCGTCCGCCACGCGTCCGAGGGACAAGAGAGAAGAGAGACTGAAACAGGGAGAAGA
GGCAGGAGAGGAGGAGGTGGGGAGAGCAGAACTGGAGGCCGACACTGAGGGAGGGCGGGA
GGAGGTGAAGAAGGAGAGAGGGGAGAAGAGGCAGGAGCTGGAAAGGAGAGAGGGAGGAGGAG
GAGGAGATGCGGGATGGAGACCTGGAGTTAGGTGGCTTGGGAGAGCTTAATGAAAAGAGAAC
GGAGAGGAGGTGTGGTTAGGAACCAAGAGGTAGCCCTGTGGGCAGCAGAAGGCTGAGAGGA
GTAGGAAGATCAGGAGCTAGAGGGAGACTGGAGGGTTCCGGGAAAAGAGCAGAGGAAAGAGG
AAAGACACAGAGAGACGGGAGAGAGAAGAAGAGTGGGTTTGAAGGGCGGATCTCAGTCCCTG
GCTGCTTTGGCATTGTGGGAACTGGGACTCCCTGTGGGGAGGAGAGGAAAGCTGGAAGTCCT
GGAGGGACAGGGTCCAGAAAGGAGGGACAGAGGAGCTGAGAGAGGGGGCAGGGCGTTGGG
CAGGGGTCCCTCGGAGGCCTCCTGGGGATGGGGGCTGCAGCTCGTCTGAGCGCCCCCTCGAGC
GCTGGTACTCTGGGCTGCACTGGGGGCAGCAGCTCACATCGGACCAGCACCTGACCCCCGAGG
ACTGGTGGAGCTACAAGGATAATCTCCAGGGAAACTTCGTGCCAGGGCCTCCTTTCTGGGGC
CTGGTGAATGCAGCGTGGAGTCTGTGTGCTGTGGGGAAGCGGCAGAGCCCCGTGGATGTGGA
GCTGAAGAGGGTTCTTTATGACCCCTTTCTGCCCCCATTAAAGGCTCAGCACTGGAGGAGAGA
AGCTCCGGGGAACCTTGTACAACACCGGCCGACATGTCTCCTTCTGCCTGCACCCCGACCT
GTGGTCAATGTGTCTGGAGGTCCCTCCTTTACAGCCACCGACTCAGTGAAC TGCGGCTGCT
GTTTTGGAGCTCGCGACGGAGCGGCTCGGAACATCAGATCAACCACAGGGCTTCTCTGCTG
AGGTGCAGCTCATTCACCTTCAACCAGGAACCTTACGGGAATTTAGCGCTGCCTCCCGCGGC
CCCAATGGCCTGGCCATTCTCAGCCTCTTTGTCAACGTTGCCAGTACCTCTAACCCATTCTT
CAGTCGCCTCCTTAACCGCGACACCATCACTCGCATCTCTACAAGAATGATGCCTACTTTC
TTCAAGACCTGAGCCTGGAGCTCCTGTTCCCTGAATCCTTCGGCTTCATCACCTATCAGGGC
TCTCTCAGCACCCCGCCCTGCTCCGAGACTGTCACCTGGATCCTCATTGACCGGGCCCTCAA
TATCACCTCCCTTCAGATGCACTCCCTGAGACTCCTGAGCCAGAATCCTCCATCTCAGATCT
TCCAGAGCCTCAGCGGTAACAGCCGGCCCTGCAGCCCTTGGCCACAGGGCACTGAGGGGC
AACAGGGACCCCGGACCCCGAGAGGCGCTCCGAGGCCCAACTACCGCCTGCATGTGGA
TGGTGTCCCCATGGTCGCTGAGACTCCCTTCGAGGATTGCACCCGCGCGTCTTAAGCCTC
CCCACAAGGCGAGGGGAGTTACCCCTAAAACAAAGCTATTAAAGGGACAGAATACTTA

FIGURE 145

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA34353

<subunit 1 of 1, 328 aa, 1 stop

<MW: 36238, pI: 9.90, NX(S/T): 3

MGAAARLSAPRALVLWAAALGAAAHIGPAPDPEDWWSYKDNLQGNFVPGPPFWGLVNAAWSLC
AVGKRQSPVDVELKRVLYDPFLPPLRLSTGGEKLRGTLYNTGRHVSFLPAPRPVVNVSGGPL
LYSHRLSELRLLFGARDGAGSEHQINHQGFSAEVQLIHFNQELYGNFSAASRGPNGLAILSL
FVNVASTSNPFLSRLLNRTITRISYKNDAYFLQDLSLELLFPESFGFITYQGSLS TPPCSE
TVTWILIDRALNITSLQMHSRLRLSQNPFSQIFQSLSGNSRPLQPLAHRALRGNRDPRHPER
RCRGPNYRLHVDGVPHGR

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 177-199

N-glycosylation site.

amino acids 118-121, 170-173 and 260-263

Eukaryotic-type carbonic anhydrases proteins

amino acids 222-270, 128-164 and 45-92

FIGURE 146

GGCGCCTGGTTCTGCGCGTACTGGCTGTACGGAGCAGGAGCAAGAGGTGCGCGCCAGCCTCCGCGCCGAGCCTC
 GTTCGTGTCCCCGCCCTCGCTCCTGCAGCTACTGCTCAGAAACGCTGGGCGGCCACCCCTGGCAGACTAAAGAA
 GCAGCTCCCTTCCACCCCAACTGCGAGGTCTAATTTTGGACGCTTTGGCTGCCATTTCTTCCAGGTTGAGGGAGC
 CGCAGAGGCGGAGGCTCGGTTATCTTCCGAGCTCAGCACCCAGCTCGCCCCGGAAGCTCGTGTCTCAGGCCCTTC
 GCGAGCGGGGCTCTCGCTCTCGGTTCCCTTGTGAAGGCTGTGGGCGGCTGCAGAGGCGGGCCGTCGGGTTTGGCT
 CACCTCTCCAGGAAACCTTCACTGTGGAGAGCAAAAGGAGTGAAGAGCCTGTCTTGGAGATTTTCTGGGGAA
 ATCCTGAGGTCACTCATTTGAAGTGTACCGCGCGGAGTGGCTCAGAGTAACACAGTGTCTGTCTATGGCTAGA
 GCAATTCAGCCATGGTGGTTCCCAATGCCACTTTATTGGAGAACTTTTGGAAAAATACATGGATGAGGATGGT
 GAGTGGTGGATAGCAAAACGAGGGGAAAAGGGCCATCAGACAGCAATGACATGCAGAGTATTTTGGACCTTCAT
 TAATAAATACGAAGTCAAGTGTATCCACAGCCTCTAATATGGAGTATATGACATGGGATGTAGAGCTGGAAAGA
 CTTCAGAAATCTGGGCTGAAAGTTGCTTGTGGGAACATGACCTGCAAGCTTGCTTCCATCAATTGGACAAAT
 TTGGGAGCACTCGGGGAAGATATAGGCCCCCGAGCTTTCATGTACAATCGTGGTATGATGAAGTGAAGACTTT
 AGCTACCCATATGAACATGAATGCAACCCATATTGTCCATTCAGGTGTTCTGGCCCTGTATGTACACATATACA
 CAGGTCGTGTGGGCACTAGTAACAGAAATCGGTTGTGCCATTAATTTGTGTACATAACATGAACATCTGGGCGCAG
 ATATGGCCCAAAGCTGTCTACCTGGTGTGCAATTACTCCCCAAAGGGAACTGGTGGGGCCATGCCCTTACAAA
 CATGGGCGGCCCTGTCTCTGTTGCCACCTAGTTTGGAGGGGGCTGTAGAGAAAAATCTGTGCTACAAAAGAGG
 TCAGACAGGTATTATCCCCCTCGAAGAGAGGAAACAAATGAAATAGAACGACAGCAGTCACAAGTCCATGACACC
 CATGTCCGGACAAGATCAGATGATAGTAGCAGAAATGAAGTCATAAGCGCACAGCAAAATGTCCCAATTTGTTCT
 TGTGAAGTAAGATTAAAGAGATCAGTGCAGAAAGGAAACAACTGCAATAGGTACGAATGTCTCTGCTGGCTGTGGAT
 AGTAAAGCTAAAGTTATGGCAGTGTACATTATGAAATGCAATCCAGCATCTGTAGAGCTGCAATTCATTATGGT
 ATAATAGACAAATGATGGTGGCTGGGTAGATATCACTAGACAAGGAAGAAAGCATTATTTCATCAAGTCCAAATAGA
 AATGGTATTCAAACAATTGGCAAAATATCAGTCTGCTAATTCCTTACAGCTCTCAAAGTAAACGTTCCAGGCTGTG
 ACTTGTGAAACAACTGTGGAACAGCTCTGTCCATTTCATAAGCCTGCTTCACATTGCCCAAGAGTATACCTGTCT
 CGTAACTGTATGCAAGCAAAATCCACATTATGCTCGTGTAAATTTGGAACCTCGAGTTTATTCTGATCTGTCCAGTATC
 TGCAGAGCAGCAGTACATGCTGGAGTGGTTCGAAATCAGCGTGGTATATGTTGATGTAAATGCTGTGGACAAAAGA
 AAGACCTACATTGCTTTCTTTCAAGATGGAATCTTCTCAGAAAGTTTACAGAAATCCTCCAGGAGGAAAGGCATTC
 AGAGTGTGTGCTGTGTGTGAACCTGAATACTTGGAGAGGAGCAATAAAGACTATTCCAAATGCAATATTTCTGA
 ATTTTGTATAAACTGTACATTACTGTACAGAGTACATCACTATTTTCAGCCCAAAAGGTCGCAAAATGCATA
 TAAATCTGTATAAAACAAAGTCTATAAAATAAAACATGGGACATTAGCTTTGGGAAAAGTAAATGAAAATATAATGG
 TTTTGAAGAATCCTGTGTAAATATTGCTATATTTCTTAGCAGTTATTTCTACAGTTAATACATAGTCATGATT
 GTTCTACGTTTCATATATTATATGGTGTCTTGTATATGCCACTAATAAAATGAATCTAAACATTGAATGTGAAT
 GCCCTCAGAAAAATCATCTAGTGCATTTAAAAATAATCGACTCTAAAACCTGAAGAGAACTTATCACATTTTCCCC
 AGTTCAATGCTATGCCATTACCACTCCAAATAATCTCAAATAATTTTCCACTTAATAACTGTAAAGTTTTTTTT
 TGTTAATTTAGGCATATAGAATAITAAATCTGATATTGCACTTTCTATTTTATATAAAATTAATCTTTTAATATC
 CAAATGAATCTGTTAAAGTGTGTGATTCCTTGGGAATGGCTTAAAAATAAATGTAGTATAAAGCTCAGAGTGGTGGT
 AGAAAAACCTTCAGTGTATCATGTAGTAAATGTAGGGTTAAACATGACAGCCAGCACTTCTATGTACTGTTA
 AAATGAGGTACATATATTTCTTTGTATCTTGGCAAATACTCCTGCGAGGAGGAGTATTAATGACAAAAGTT
 GCAACAAAGTCAACTAATGTATTACATTACCATTGCACTGATTTTAAATGTGAAATGACCTGTGTATATAA
 ATATTGCCATATCATGTACCTATAATGGTGATATATTGTTTCTATGAAAATGTATTGTGCTTTGATACTAAA
 AATCTGTAAATGTACTTTTGGTAATTTTCTTCTGCTGGTGGATTACATATTAAATTTTCTGCTGGTGGG
 TAAACATTAAATTAATCATGTTTCAAAAAA

FIGURE 147

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45417

<subunit 1 of 1, 500 aa, 1 stop

<MW: 56888, pI: 8.53, NX(S/T): 2

MKCTAREWLRVTTVLFMARAIPAMVVPNATLLEKLLEKYMDEEDGEWWIAKQRGKRAITDNDM
QSILDLHNKLR SQVYPTASNMEYMTWDVELERSAESWAESCLWEHGPA SLLPSIGQNLGAHW
GRYRPPTFHVQSWYDEVKDFSYPYEHECNPYCPFRCSGPVCTHYTQV VWATSNRIGCAINLC
HNMNIWGQIWPKAVYLV CNYSPKGNWWGHAPYKHGRPC SACPSPFGGCRENL CYKEGSDRY
YPPREEETNEIERQQSQVHDTHVTRSDSSRNEVISAQQMSQIVSCEVRLRDQCKGTT CNR
YBCPAGCLDSKAKVIGSVHYEMQSSICRAAIHYGII DNDGGVVDITRQGRKHYFIKSNRNGI
QTIGKYQSANSFTVSKVTQAVTCETTVEQLCFPHKPASHCPRVYCP RNCMQANPHYARVIG
TRVYSDLSSICRAAVHAGVVRNHGGYVDVMPVDKRKTYIASFQNGIFSES LQNPPGGKAFRV
FAVV

Important features:

Signal peptide:

amino acids 1-20

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 protein

amino acids 165-186, 196-218, 134-146, 96-108 and 58-77

N-glycosylation site

amino acids 28-31

FIGURE 148

GCGGAGACAAGCGCAGAGCGCAGCGCACGGCCACAGACAGCCCTGGGCATCCACCGACGGCG
CAGCCGGAGCCAGCAGAGCCGGAAGGCGCGCCCCGGGCAGAGAAAGCCGAGCAGAGCTGGGT
GGCGTCTCCGGGCGCGCTCCGACGGGCCAGCGCCTCCCCATGTCCCTGCTCCCACGCCG
CGCCCTCCGGTCAGCATGAGGCTCCTGGCGGCCGCGCTGCTCTGTGCTGCTGGCGCTGT
ACACCGCGCGTGTGGACGGGTCCAAATGCAAGTGCTCCCGGAAGGGACCCAAGATCCGCTAC
AGCGACGTGAAGAAGCTGAAATGAAGCCAAAGTACCCGACTGCGAGGAGAAGATGGTTAT
CATCACCACCAAGAGCGTGTCCAGGTACCGAGGTCAGGAGCACTGCCTGCACCCCAGCTGC
AGAGACCAAGCGCTTCATCAAGTGGTACAACGCCTGGAACGAGAAGCGCAGGGTCTACGAA
GAATAGGGTGAAAAACCTCAGAAGGGAAAACTCCAAACCAGTTGGGAGACTTGTGCAAAGGA
CTTTGCAGATTAAAAAAGCTTTTCTCACAGGCATAAGACACAAATTATATATTGTTATGAAGCACTTTTACCAACGGTCAG
TTTTTACATTTTTATAGCTGCGTGCAGAAAGGCTTCCAGATGGGAGACCCATCTCTCTTGTGCT
CCAGACTTCATCACAGGCTGCTTTTTATCAAAAAGGGGAAAACTCATGCCTTTCCTTTTTAA
AAAATGCTTTTTTGTATTGTGCCATACGTCACTATACATCTGAGCTTTATAAGCGCCCCGGA
GGAACAATGAGCTTGGTGGACACATTTCAATTGCAGTGTGCTCCATTCCTAGCTTGGGAAGC
TTCGCTTAGAGGTCTGGCGCCTCGGCACAGCTGCCACGGGCTCTCCTGGGCTTATGGCCG
GTCACAGCCTCAGTGTGACTCCACAGTGGCCCCGTAGCCGGGCAAGCAGGAGCAGGTCTCT
CTGCATCTGTTCTCTGAGGAACCTCAAGTTTGTTGCTGCCAGAAAAATGTCTTCAATTCCCCCT
GGTTAATTTTTACACACCCCTAGGAAACATTTCCAAGATCCTGTGATGGCGAGACAAATGATC
CTTAAAGAAGGTGTGGGGTCTTTCCCAACCTGAGGATTTCTGAAGAGTTACAGGTTCAATA
TTTAATGCTTCAGAAGCATGTGAGGTTCCCAACACTGTCAGCAAAAACCTTAGGAGAAAAC
TAAAAATATATGAATACATGCGCAATACACAGCTACAGACACACATTCTGTTGACAAGGGAA
AACCTTCAAAGCATGTTTCTTTCCCTCACCACAACAGAACATGCAGTACTAAAGCAATATAT
TTGTGATTCCCATGTAAATCTTCAATGTTAAACAGTGCAGTCTCTTTTCGAAAGCTAAGAT
GACCATGCGCCCTTTCTCTGTACATATACCCTTAAGAACGCCCCCTCCACACACTGCCCCC
CAGTATATGCCGATTGTACTGCTGTGTTATATGCTATGTACATGTACAGAAACCATTAGCAT
TGCATGCAGGTTTCATATTCTTTCTAAGATGGAAAGTAATAAAATATATTTGAAATGTAAAA
AAAAAAAAAA

FIGURE 149

MSLLPRRAPPVSMRLLAAALLLLLLALYTARVDGSKCKCSRKGPKIRYSDVKKLEMKPKYPH
CEEKMVIITTKSVSRYRGQEHCLHPKLQSTKRFIKWYNAWNNEKRRVYEE

Signal sequence:

amino acids 1-34

100
90
80
70
60
50
40
30
20
10
0
100
90
80
70
60
50
40
30
20
10
0

FIGURE 151

MLNSNVLLWLTALAIKFTLIDSQAQYPVVNTNYGKIRGLRTPLPNEILGPVEQYLGVPYASP
PTGERRFQPEPPSSWTGIRNTTQFAAVCPQHLDERSLLHDMLPIWFTANLDTLMTYVQDQN
EDCLYLNIIYVPTEDGANTKKNADDITSNDRGEDEDIHDQNSKKPVMVYIHGGSYMEGTGNMI
DGSILASYGNVIVITINYRLGILGFLSTGDQAAKGN YGLLDQIQALRWIEENVGAFGGDPKR
VTIFGSGAGASCVSLLTLSHYSEGLFQKAI IQSGTALSSWAVNYQPAKYTRI LADKVGCNML
DTTDMVECLRNKNYKELIQQTITPATYHIAFGPVIDGDVIPPDDPQILMEQGEFLNYD IMLGV
NQGEGLKFVDGIVDNEDGVTPNDFDFSNSFVDNLYGYPEGKDTLRETIKFMYTDWADKENP
ETRRKTLVALFTDHQWVAPAVAADLHAQYGSPTYFYAFYHHCQSEMKP SWADSAHGDDEVPIV
FGIPMIGPTELFSCNFSKNDVMSAVVMTYWTNFAKTGDPNPQVPQDTKFIHTKPNRFEEVA
WSKYNPKDQLYLHIGLKPRVRDHYRATKVAFWLELVPHLHNLNEIFQYVSTTTKVP PPDMS
FPYGTRRSPAKIWPTTKRPAITPANPNKHSKDPHKTGPEDTTVLIETKRDYSTELSVTI AVG
ASLLFLNILAFAALYYKDKRRRHETHRRPSPQRNTTNDIAHIQNEEIMSLQMKQLEHDHECE
SLQAHDTLRLTCPPDYTLTLRRSPDDIPLMTPNTITMIPNTLTGMQPLHTFNTFSGGQNSTN
LPHGHSTTRV

Signal sequence:

amino acids 1-24

Transmembrane domains:

amino acids 189-204, 675-692

FIGURE 152

GGGAAAGATGCGCGGCAGCTCTGGGACCCCTTGGGTCGTGGCAGCAGTGGCGGCGATGTTTGT
CGGCTCGGGATGGGTCCAGGATGTTACTCCTTCTTCTTTGTTGGGGTCTGGGCAGGGGCCA
CAGCAAGTCGGGCGGGTCAAACGTTTCGAGTACTTGAACGGGAGCACTCGCTGTGCAAGCC
CTAGCAGAGTGTGGGCACAGGCAGTTTCCTCACTGTGGAACTGTGATGGGCATGCCCATTGGTGA
TGACCCAGTATATCCCGCTTACCOCAGATATGCAAGTAACAGGGTGCCCTTGTGGAAACCGG
GTGCCATGTTTTCTGAGACCTGGGAGTTGCAGGTGCACCTTCAAATCCATGGACAAGGAAA
GAAGAATCTGCATGGGGATGGCTTGGCAATCTGGTACACAAAGGATCGGATGCAGCCAGGGC
CTGTGTTTGGAAACATGGACAAATTTGTGGGGCTGGGAGTATTTGTAGACACCTACCCCAAT
GAGGAGAAGCAGCAAGAGCGGGTATTCCCTACATCTCAGCCATTGGTGAACAAAGCTCCCT
CAGCTATGATCATGAGCGGGATGGGCGGCCACAGAGCTGGGAGGCTGCACAGCCATTGTCC
GCAATCTTCATTACGACACCTTCCCTGGTGATTTCGCTACGTCAAGAGGCATTGTGACGATAATG
ATGGATATTGATGGCAAGCATGAGTGGAGGGACTGCATTGAAGTGCCCGAGTCCGCCTGCC
CCGCGGCTACTACTTCGGCACCTCCTCCATCACTGGGATCTCTCAGATAATCATGATGTCA
TTTCCTTGAAGTTGTTTGAATGACAGTGGAGAGAACCACGAGGAAAAGCTCCATGCGA
GATGTGTTCTTGGCCCTCAGTGGACAATATGAAGCTGCCTGAGATGACAGCTCCACTGCCGCC
CCTGAGTGGCTTGGCCCTCTTCCCTCATCGTCTTTTCTCCCTGGTGTGTTTCTGTATTGTGCA
TAGTCATTGGTATCATACTCTACAACAAATGGCAGGAACAGAGCCGAAAGCGCTTCTACTGA
GCCCTCCTGCTGCCACCACCTTTGTGACTGTCAACCATGAGGTATGGAAGGAGCAGGCACTG
GCCTGAGCATGCAGCCTGGAGAGTGTCTTGTCTCTAGCAGCTGGTTGGGGACTATATTCTG
TCACTGGAGTTTTGAATGCAGGGACCCGCAATCCATGTTGTGTCATGGGGACATCAACT
CTGGTCTGGGAAGCCACCCACCCAGGGCAATGCTGCTGTGATGTGCCCTTTCCTGTCAGTCC
TTCATGTGGGAGCAGAGGTGTGAAGAGAATTTACGTGGTTGTGATGCCAAATACAGAAAC
AGAATTTCATAGCCCAGGCTGCGGTGTTGTTTGACTCAGAAAGGCCCTTCTACTTCAGTTTTG
AATCCACAAGAATTAATAAATCTGGTAACACCAGAGCTTTCTGACCATCCATTTCGTTGGGTT
TTGCATTTGACCCAACCTCTGCCTACCTGAGGAGCTTCTTTGGAAACAGGATGGAAGT
TCTTCCCTGCCCTTACCTTCCCTTCACTCCATTATTGCTCTCTGTGTGCAACCTGAGCTG
GGAAAGGCATTGGATGCCCTCTCTGTTGGGGCCTGGGGCTGCAGAACACACCTGCGCTTTCAC
TGGCCTTCATTAGGTGGCCCTAGGGAGATGGCTTTCGCTTTGGATCACTGTTCCCTAGCAT
GGGTCTTGGGTCTATTGGCATGTCCATGGCCTTCCCAATCAAGTCTCTTCAGGCCCTCAGTG
AAGTTTGGCTAAAGGTTGGTGTAAAAATCAAGAGAAGCCTGGAAGACATCATGATGCCATG
GATTAGCTGTGCAACTGACCAGCTCCAGGTTTGTATCAAACCAAAGCAACATTGTGTCATGTG
GTCTGACCATGTGGAGATGTTTCTGGACTTGTAGAGCCTGCTTAGCTGCATGTTTGTAGT
TACGATTTTGGAAATCCCACTTTGAGTGTCTGAAAGTGTAAAGGAAGCTTCTTCTTACACCTT
GGGCTTGGATATTGCCAGAGAAGAAATTTGGCTTTTTTTTTTCTTAATGGACAAGAGACAGT
TGCTGTTCTCATGTTCCAAGTCTGAGAGCAACAGACCCTCATCATCTGTGCCCTGGAAGAGTT
CACTGTCAATTGAGCAGCACAGCCTGAGTGTGCCTCTGTCAACCTTATCCACTGCCTTA
TTTGACAAGGGGTACATGCTGCTCACCTTACTGCCCTGGGATTAATACAGTTACAGGCCAG
AGTCTCCTTGGAGGCGCTGGAACTCTGAGTCTCCTATGAACCTCTGTAGCCTTAATGAAT
TCTTAAATACCCGATGGAACCAAAAAAAAAAAAAAAAAAGGGCGGCCGCACTCTAGAGTGC
ACCTGCAGTAGGGATAACAGGGTAATAAGCTTGGCCGCCATGG

FIGURE 153

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50911
><subunit 1 of 1, 348 aa, 1 stop
><MW: 39711, pI: 8.70, NX(S/T): 1
MAATLGPLGSWQQWRRCLSDGSRMLLLLLLLGSGQGPPQVAGQTFEYLKREHSLSKPYQ
GVGTGSSSLWNLMGNAMVMTQYIRLTPDMQSKQGALWNRVPCFLRDWELQVHFKEHGQGKK
LHGDGLAIWYTKDRMQPGPVFGNMDKFVGLGVFVDTYPNEEKQQERVFPYISAMVNNGSLSY
DHERDGRPTELGGCTAIVRNLYHDTFLVIRYVVRHLTIMMDIDGKHEWRDCIEVPGVRLPRG
YYFGTSSITGDLSDNHDVISLKLFEFTVERTPEEEKLHRDVFLPSVDNMKLPMTAPLPPLS
GLALFLIVFFSLVFSVFAIVIGIILYNKWQEQSRKRFY
```

Signal sequence:

amino acids 1-38

Transmembrane domain:

amino acids 310-329

FIGURE 154

CCGAGCCGGGCGCGCAGCGACGGAGCTGGGGCCGGCCTGGGACCATTGGGCGTGAGTGCAATCTACGGATCAGTCT
 CTGATGGTGGGTGCTTAAACCTCAGTGGGGACTCCAGGATTTCCATGAAGAAAAATCAGTTGTCTTTCATTCAAGAAT
 TGGGGTCTGGCTCAGAATTCTCTGAGCTGGTGAAAACTCTGTTTTCTAGAAGAGGTTTAAITTAATGCTCGAGTCT
 GACATGTTCCCGATTTGAGGTGAAACCATGAAGAGAAAAAGAACTACTTAAATAATGCTTTTCGCGAACCGCTTCT
 TGCTGCTGCTGGCCCTGGCTCGCTGCTGGCCCTTTGTGAGCCTCAGCCTGCGAGTTCTCCACCTGATCCCGGTGT
 CGACTCTTAAAGAAATGGAATGAGTAGCAAGAGTCGAAAGAGAATCATGCCCCGACCTGTGACGGAGGCCCTGTGA
 CAGACCCCGTTTATGAAGCTCTTTTGTACTGCAACATCCCCAGTGTGGCCGAGCGCAGCATGGAAGGTCTATGCC
 CGCATCATTTTAAAGCTGGTCTCAGTGCAATGTTTCATTCGCCACGGAGACAGGTACCCACTGTATGTCTATCCCA
 AACAAGAGCGACAGAAATTTGACTGCACTCTGGTGGCTAACAGGAAACCGTATCACCAAAACCTGGAAGCTTTTCA
 TTAGTGCATGTCGAAAGGATCCGGAGCCTCTTTGAAAGCCCTTTGAAGCTCTTGGCTCTTTTACCCAAATCACC
 CATTGTGTGAGATGGGAGAGCTCAACAGACAGGAGTTGTGCGAGCTTTGCAAGACCGTCACTGCTGAGGAGATA
 TCTATCTAAAGAAACCAAACTCTCTGCCAATGATGGTCTGCGACACAGCTCTATTTAGAGACCCTGGGAAAA
 GCGCGACCTTCAAAAGTGGCTGGCCTTGCTTTATGGCTTTCTCCAGATTTTGACTGGAAGAGATTTTATTTCA
 GGCACCGACCAAGTGGCTGTCTTGCTCTGGAAGCTGCTATTGCCCCGTGAAGAAACAGTATCTGGAAGAGGAGC
 AGCGTCTGCATACCTCTCTAGTTTGAAGAACAGCCAGCTGGGAGAGACCTACGGGGAGATGGCCAGATGCTGGG
 ATGTCCCCACCAAGCAGCTTAGAGCTGCCAACCCCATAGACTCCATGCTCTGCCACTTCTGCCCAATGTCAGT
 TTCCTGTACAGAAATGGCTGTGTGACATGGAGCACTTCAAGGTAAATAGAGCCCATCAGATCGAGGATGAAA
 GGGAAAGACGGGAGAGAAGAAATTTGACTTCGGGTATTTCTTCTGGGTGCCCAACCCATCTCGAACCACAAACCATCG
 GCGGATGACGCTGCCACCGGAGGCGAGGAAAGAGGCTCTTGGCCCTCTACTGTGCTCATGATGTCACTCTGT
 CACCGATTCTCAGTGCTTTGGGCTTTTCAAGCCAGGTTCCCAAGGTTTGACGAGCCAGGTTGATCTTTGAGCTTT
 GGCAGACAGAGAAAAGCCAGTGAACATTCCTGCCGATTTCTTACAATGGGCTCGATGTCACATTTCCACACT
 CTTTCTGCGAAGACCCACCAAGGCTTCTCCCAAGCCATGTGCGCCCTTGAAAACTTTGGTCCGCTTTTGTGAAAA
 GGGACATGTTGTAGCCCTGGGTGGCAGTGATCAAAATATTATGATGCAATGTCACAGGGAAGGATTTAAAGG
 TATGCACTACAGCAGTATAGAAATCCATGCCAATACAGAGCATAGGGAAAGGTCCACTTCTAGTTTGTGCTGTTAC
 TAAAGGTAGAAGATTTATGCTTTTAAAGGCTAAATATGTTTGTGGGAACCAAGATGGTGGGGTTGAACAGT
 AAGCATTGTCTGCAATGTGGTACGTGAATTTGCTTGGTACAAAATGGCCAGTTTACAGAGGAATAGAAGGTACTT
 TATCATAGCCAGACTTCGCTTAGAATGCCAGAAATATATAGTTCAAGACCTGAAGTTGCCAATCCAGTTTGCAC
 TCTTCTGGCCTGCCCATGTTACTATGTGATGGAACAGCACACCTCAACAAAAATTTTTTAATCTTAGACATT
 TTTACCTTGTCTTGTGAAGAAATTTCTGAAGTGATTTATCTAAAAATAAGGTTGGCAAACTTTTCTGTGAAGG
 GCCAGATTGTAATATTTTACAGCTGTGTGGACAAAAGGCCACATACAGTCTCTGTGCTAACTACTCAACTCTGT
 TTTGAAGCAGGAAGCCACACAGACAGTACATAAAGGAATATGTGTAGCTGGGTCCCGGCCAGACAAAAACA
 GATGGTGACCAGACTTGGCCCTGGGCTGTAGTTTGTGACCCCTCATCTAAAAAATAGGCTATACATACTAGAA
 ACTTCCAGCACTTTGAGACAGAGTTGAATACCAAGAAATATTCAATGGTTCTCCAGTAACTTCTGCTAGAAAA
 CAGAAATTTGGTCTGATCTGACACTAGAACAAAACTTTGAGGGTAAATAAACATTTGAATAGAATGAATCATAGAA
 AACTGATTAGAAGAACTTTGATGTTTATGATGATTTGGTACAAGATAGTTTAAAGTATGTTCTAAATATTTGT
 CTGTGTAGTCTATTGCTGTATATGCTGAAATTTTGTATGCACTTTAGTATTTTATAGTTTGAAGAAATATT
 TTTCAAGACAGTTTATGATGACTCTTTATCTGTAGTAAATATTCAATTTGCTGTGACTGCTGGTGGTGAAG
 GAGGCTAGAAGATGAATCAGGCACCTTTCTCAATAAAACCTAATATGAGCTTATCCCTTTGACAGCTGTGAGA
 ACTGGATTCATTTTAAACCAATTTTATCAGTTTCAATGTAATCTGATTGATTTTAAATGCGTTTGTGA
 AGAAGCTTGTCTATTAGGTATTTTACAGATCTTTATAAGGTGTTTATATATAGAAGCAATTTAATATACATCTG
 TGAATTTCTGAACATAATGSGTCAATTCAGAGAAATGGAAGTGAAGTGAAGTCTGTCTGTCTCGCATTC
 AACTTTTTCTCTTTGTTTTGTCCAGTGTGCAATTTGAATATGCTCTTTCTATATAATAAATTTTTTAAGAATA

FIGURE 155

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48329

><subunit 1 of 1, 480 aa, 1 stop

><MW: 55240, pI: 9.30, NX(S/T): 2

MLFRNRFLLLALLAALLAFVSLSLQFFHLIPVSTPKNGMSSKSRKRIMPDVTEPPVTPDPVY
EALLYCNIPPSVAERSMEGHAPHHFKLVSVHVFIRHGDYPLYVIPKTKRPEIDCTLVANRKP
YHPKLEAFISHMSKSGSGASFESPLNSLPLYPNHPLCEMGELTQTGVVQHLQNGQLLRDIYLK
KHKLLPNDWSADQLYLETTGKSRTLQSGLALLYGFLPDFDWKKIYFRHQPSALFCSGSCYCF
VRNQYLEKEQRRQYLLRLKNSQLEKTYGEMAKIVDVPTKQLRAANPIDSM LCHFC HNVSFPC
TRNGCVDMEHFVKVIKTHQIEDERERREKKLYFGYSLLGAHPILNQTIGRMQRATEGRKEELF
ALYSAHDVTLSPLSALGLSEARFPRFAARLIFELWQDREKPSEHSVRILYNGVDVTFHTSF
CQDHHKRSPPKMCPLENLVRFVKRDMFVALGSGTNYIDACHREGF

Signal sequence:

amino acids 1-18

FIGURE 156

AAAAAAGCTCACTAAAGTTTCTATTAGAGCGAATACGGTAGATTTCATCCCCCTTTTGAAGAAACAGTACTGTGGA
 GCTATTTAAGAGATAAAAAACGAAATATCCTTTCTGGGAGTTCAGATTGTGCAGTAAATTTGGTTAGGACTCTGAGC
 GCCGCTGTTTACCAATCGGGGAGAGAAAAACGGAGATCCTGCTCGCCTTGACGCGCCTTGAAGCACAAGACGAT
 AGCTAGGAATGAACCATCCCTCGGGAGTATGTGGAACAACAGGAGGAGCTCTGACTTCCCACTGTCCCATTTCTAT
 GGGCGAAGGAACCTGCTCTGACTCTCAGTGGTTAAGGGCAGAAATGAAAAATAATCTGGAGGAAAGATAAAGATGAT
 TCTTGCGGAGCTGCACCGGACTACAAGGGCTTGTCTGCTGGGAATCCTCCTGGGGACTCTGTGGGAGACCGG
 ATGCACCCAGATACGCTATTTCAGTTCCGGAAGAGCTGGAGAAAGGCTCTAGGGTGGGCGACATCTCCAGGGACT
 GGGGCTGGAGCCCGGGAGCTCGCGGAGCGCGAGTCCGCATCATCCCAGAGGTAGGACCGAGCTTTTCGCCCT
 GAATCGCGCGAGCGGACGCTTGGTCACGGCGGGCAGGATAGACGGGAGGAGCTCTGTATGGGGCCATCAAGTG
 TCAATTAATCTAGACATCTGATGGAGGATAAAGTAAAAATATATGGAGTAGAAGTAGAAGTAAAGGGACATTAAG
 CGACAATGGCGCTTACTTTTCGTAAAGTGAATTAGAAAAATAAATAGTGAACCACTAGGACCACTGAGATGGCGGTT
 CCCTCTACCCACGCCCTGGGATCGGATATCGGGAAGAACTCTCTGCAGAGCTACGAGCTCAGCCCBAACACTCA
 CTTCTCCCTCATCGTGCAAAATGGAGCCGACGCTAGTAGTAAGTACCCGAATTTGGTGCTGAAACGCGCCCTGGACCG
 CGAAGAAAAGCTGCTCACCACTGGTCTTACGCGCTCCGACGGGGCGACCCGCTGCGACAGGACACCGCGCG
 CATCCGCTGATGGTTCTGGGACGACGAGCTGCTTGTAGTCAACGCTACCGACCTGACGAAGGAGTCAATGGGAAGT
 GAGGATATCTCTCCGGTATGTGGACGACAGGCGCGCCCAAGTTTTCAACTAGATTGTAATTCAGGGACAATATC
 AACAAATAGGGGAGTTGGACCAAGAGGAGTCAGGATTCTACCAGATGGAAGTGAAGCAATGGATAATGCAGGATA
 TTTCTGCGGAGCCAAAGTCTGATCACTGTTCTGACGTGAACGACAATGCCCCAGAAGTGGTCTCACTCTCTCT
 CGCCACTCGGTTCCCGAAAACTCTCCAGAGGGACATTAATTGCCCTTTAAATGTAAATGACCAAGATTCTGA
 GGAACACGACAGGTGATCTGTTTCATCCAGGAAATCTGCCCTTTAAATTAGAAAAATCTTACGGAATTACTA
 TAGTTTAGTACAGACATAGTCTTGGATAGGGAAACAGGTTCTAGCTACAAACATCAAGTGAACCGCCACTGACCG
 GGGAAACCCGCCCCCTATCCACGGAACCTCATATCTGCTGAACGTGGCAGACACCAAGCAACCCCGCGGTCTTT
 CCCTCAGGCTCCTTATTCCTGTTATATCCAGAGAAACATCCAGAGGAGTTCCTCGCTCTGTGACCCGCCA
 CGACCCCGACTGTGAAGAGAACGCCAGATCACTATTCTCCCTGGCTGAGAACACCATCCAGGGGGCAAGCCTATC
 GTCTCACTGGTCCATCAAGTCCGACACTGGGGTACTGTATGCGCTGAGCTCCTTCGACTACGACGAGTTCCGAGA
 CTTGCAAGTGAAGTGATGGCGGGGACACGGGACCCGCCCTCAGCAGCAACGTTGCTGTGAGCCTGTTCTGT
 GTGGACAGAACGACAATGCGCCCGAGATCCTGTACCCCGCCTCCCAACGGACGGTTCCACTGGCGTGGAGT
 GGTCTCCCGCTCCGACAGGCCCGGCTACCTGTTGACCAAGGTGGTGGCGGTGGAACAGACTCTCGCGCAGAAGCC
 CTGCTGCTCACTACGCTCTGCTCAAGGCCAGCGAGCCGGACTCTTCTCGGTGGGTCTGCACACGGGCGAGGTGCG
 CACGGCGGAGCCCTGCTGGAACAGAGCGCGCTCAAGCAGAGCCTCTGATGGCGCTCAGGACACCGGCCAGCC
 CCCTCTCTCCGCACTGTCAAGCTCACCGTGGCGTGGCCGACAGCATCCCCAAGTCTGGCGGACCTCGGCGAG
 CCTCGAGTCTCCAGCTAACTCTGAACCTCAGACCTCACTCTGTACCTGGTGGTAGCGGTGGCGCGGTCTCCTG
 CGTCTTCTTGGCCTTCTGCTATCTTGTGCTGGCGCTCAGGCTGCGCGCTGGACAAGTCTGCTGACGCTGACGCG
 TTCAGGAGCGGCTTGACAGGAGCGCGCGCTCGCACTTTGTGGCGTGGACGGGCTGAGCGGCTTCTTCGACAGC
 CTATTCACAGAGTTTCCTCACCACGGAATCTGCGGAAGAGTCACTCTGTCCCGCCCAACTGATGCGAGCA
 GCTGCTCGCTCAGCGAGGAGCTTTGAAAAAGCGAGCCCTTTTGTCTGAGGTGATTCGGTATTTCTTAAGA
 CAGTCAATGAACAGGACTACAGTGGAGTGGTTATATCAAACTCTTCTTCTTTTATTAATGGCTCTGCTTCCCAAGC
 TGGAGTGAACGGGTGACGATCACTACTCTCACTGCGGCTCAAACTCTAGGCTCAAGCAATTAATCCCACTTTGCTC
 CGGGTGAACAGGACTACAGTGGTGCACGCACTACTGTCTGCTATCTATCTATCTATCTATCTATCTATCTAT
 CTATCTATCTATCTATCTATCTTCTTCTGTACAGAGGGAGTCTCAGCCTGTAATCCAGTACTTTGGGAGGC
 CGAGGCGGTGGATTCACTTGAGTGGAGTTGGAGTTGAGACAGCCGCAACATGAGAGAAACCCGCTCTATACTAA
 AAAAAATACAAATAGTCCGGGCGTGGTGGTGCATGTCTGTAAATCCAGCTATGTGGGAGGCTGAGTCAGGAGAT
 TGCTTTAACTGGGAGGTGGAGGTGGCAATGAGCTGAGATTGTGCCATTGCACCTCAGCCTGGGCAACAGAGTG
 AAACCTCTATCTCA

FIGURE 157

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48306

><subunit 1 of 1, 916 aa, 1 stop

><MW: 100204, pI: 4.92, NX(S/T): 4

MIPARLHRDYKGLVLLGILLGLTWETGCTQIRYSVPPELEKSGRVGDISRDLGLEPRELAER
GVRIIPRGRTQLFALNPRSGSLVTAGRIDREELCMGAIKCQLNLDILMEDKVKIYGVEVEVR
DINDNAPYFRESELEIKI SENAATEMRFLPHAWDPDIGKNSLQSYELSPNTHFSLIVQNGA
DGSKYPELVLKRALDREEKAAHHLVLTASDGGDPVRTGTARIRVMVLDANDNAPAFQAQPEYR
ASVPENLALGTQLLVVNATDPDEGVNAEVRYSFRYVDDKAAQVFKLDCNSGTISTIGELDHE
ESGFYQMEVQAMDNAGYSARAKVLI TVLDVNDNAPEVVLTSLASSVPENSPRGTLIALLNVN
DQDSEENGQVICFIQGNLPFKLEKSYGNYISLVTDIVLDREQVPSYNITVTATDRGTPPLST
ETHISLNVADTNDNPPFPQASYSAYIPENNPRGVSLVSVTAHDPDCEENAQITYSLAENTI
QGASLSSYSVINS DTVGLYALSSFDYEQFRDLQVKVMARDNGHPPPLSSNVSLSLFVLDQNDN
APEILYPALPTDGSTGVELAPRSAEPGYLVTKVVAVDRD SGQNAWLSYRLLKASEPGLFSVG
LHTGEVTRARALLDRDALKQSLVVAVQDHGQPPLSATVTLTVAVADSIPQVLADLGSLES PA
NSETSDLTLYLVVAVAAVSCVFLAFVILLALRLRRWHKSRLQASGGGLTGAPASHFVGVD
GVQAFQTYSHFVSLTTDSRKSHLIFPQPNYADMLVQESFEKSEPLLLSGDSVFSKDSHGL
IEVSLYQIIFLFFNCVSQAGVQRYDHSSSLRPQTPRLKQLSHLCLRCNRD YRCKPPTVCL S
IYLSIYLSIYLSIYLLSCTDGS LTPVIPVLWEAEAGGSPEVGSLRPA

Signal sequence:

amino acids 1-30

Transmembrane domains:

amino acids 693-711, 809-823, 869-888

FIGURE 158

CCCAGGCTCTAGTGCAGGAGGAGAAGGAGGAGGAGCAGGAGGTGGAGATTCCAGTTAAAAG
GCTCCAGAATCGTGTACCAGGCAGAGAACTGAAGTACTGGGGCCTCCTCCACTGGGTCCGAA
TCAGTAGGTGACCCCGCCCTGGATTCTGGAAGACCTCACCATGGGACGCCCCGACCTCGT
GCGGCCAAGACGTGGATGTTCTTGCTGGGGGAGCCTGGGCAGGACACTCCAGGGC
ACAGGAGGACAAGGTGCTGGGGGTCTGAGTGCCAACCCATTGCGAGCCTTGGCAGGCGG
CCTTGTTCCAGGGCCAGCAACTACTCTGTGGCGGTGTCCTTGAGGTGGCAACTGGGTCTT
ACAGCTGCCCACTGTAAAAAACCGAAATACACAGTACGCCTGGGAGACCACAGCCTACAGAA
TAAAGATGGCCAGAGCAAGAAATACCTGTGGTTTCAGTCCATCCACACCCCTGCTACAACA
GCAGCGATGTGGAGGACCACAACCATGATCTGATGCTTCTTCAACTGCGTGACCAGGCATCC
CTGGGGTCCAAAGTGAAGCCCATCAGCCTGGCAGATCATTGCACCCAGCCTGGCCAGAAGTG
CACCGTCTCAGGCTGGGGCACTGTACCAGTCCCCGAGAGAATTTTCTTGACACTCTCAACT
GTGCAGAAGTAAAAATCTTTCCCAGAGAAGTGTGAGGATGCTTACCGGGGAGATCACA
GATGGCATGGTCTGTGCAGGCAGCAGCAAAGGGGCTGACACGTGCCAGGGCGATTCTGGAGG
CCCCCTGGTGTGTGATGGTGCATCCAGGGCATCACATCTGGGGCTCAGACCCCTGTGGGA
GGTCCGACAAACCTGGCGTCTATACCAACATCTGCCGTACCTGGACTGGATCAAGAAGATC
ATAGGCAGCAAGGGCTGATTCTAGGATAAGCACTAGATCTCCCTTAATAAACTCACAACTCT
CTGGTTC

FIGURE 159

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48336
<subunit 1 of 1, 260 aa, 1 stop
<MW: 28048, pI: 7.87, NX(S/T): 1
MGRPRPRAAKTWMFLLLLGGAWAGHSRAQEDKVLGGHECQPHSQPWQAALFQQQQLLCGGVL
VGGNWWLTAAHCKKPKYTVRLGDHSLQNKDGPEQEI PVVQSI PHPCYNSSDVEDHNDLMLL
QLRDQASLGSKVKPISLADHCTQPGQKCTVSGWGTVTSPRENF PDTLNCAEVKIFPQKKCED
AYPGQITDGMVCAGSSKGADTCQGDSGGPLVCDGALQGITSWGS DPCGRSDKPGVYTNICRY
LDWIKKIIGSKG

Important Features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 51-71

N-glycosylation site.

amino acids 110-113

Serine proteases, trypsin family, histidine active site.

amino acids 69-74 and 207-217

Tyrosine kinase phosphorylation site.

amino acids 182-188

Kringle domain proteins motif

amino acids 205-217

FIGURE 160

GGCGCCGGTGCACCGGGCGGGCTGAGCGCCTCCTGCGGCCCGGCTGCGCGCCCGGCCCGC
CGCGCCGCCACGCCCCAACCCCGGCCCGCGCCCCCTAGCCCCCGCGCCCGGCCCGCGCCCGC
GCCGCGCCAGGTGAGCGTCCGCCCGCGCGAGGCCCGGCCCGCCCGCCCGCCCGC
CCCCGGCCGGCGGGGAACCGGGCGGATTCTCGCGCGTCAAACCACTGATCCCCATAAAAC
ATTTCATCTCCCGGCCCGCGCTGCGAGCGCCCCCGCCAGTCCGCGCGCCCGCCCGCCCTCG
CCCTGTGCGCCCTGCGCGCCCTGCGCACCCGCGGCCCGAGCCAGGCCAGAGCCGGCGGGAGC
GGAGCGCGCCGAGCCTCGTCCCGCGCCGGGCCGGGGCCGGGCGTAGCGCGCGCCCTGGA
TGGGACCCCGGCCCGGGGAGACGGGGCGCCCGCCCCGAACGACTTTCAGTCCCCGACGCGC
CCGCCCCAACCCCTACGATGAGAGGGCGTCCGCTGGAGGGAGCCGGCTGCTGGCATGGGTG
CTGTGGCTGCGAGCCTGGCAGGTGGCAGCCCCATGCCAGGTGCTTGGTATGCTACAATGA
GCCCAAGGTGACGACAAGCTGCCCCAGCAGGGCTGCAAGGCTGTGCCGTGGGCATCCCTG
CTGCCAGCCAGCGCATCTTCTGTCACGGCAACCGCATCTCGCATGTGCCAGCTGCCAGCTTC
CGTGCCCTGCGGCAACCTCACCATCTGTGGCTGCACTCGAATGTGTGGCCGAATGTATGTC
GGCTGCCTTCACTGGCTGGCCCTCCTGGAGCAGCTGGACCTCAGCGATAATGCACAGCTCC
GGTCTGTGGACCTGCCACATTCCACGGCCTGGGCCGCTACACACGCTGCACCTGGACCGC
TGCGGCCCTGCAGGAGCTGGGCCCGGGGCTGTTCGCGCGCTGGCTGCCCTGCAGTACCTCTA
CCTGCAGGACAACGCGCTGCAGGCACTGCCTGATGACACCTTCGCGACCTGGGCAACCTCA
CACACCTTCTCTGCAACGGCAACCGCATCTCCAGCGTGCCCGAGCGCGCCTTCCTGGGGCTG
CACAGCCTCGACCGTCTCTACTGCACAGAACCGCGTGGCCCATGTGCACCCCGCATGCCCTT
CCGTGACCTTGGCCGCTCATGACACTCTATCTGTTTGCCAAACATCTATCAGCGCTGCCCA
CTGAGGCCCTGGCCCCCTGCGTGCCCTGCAGTACCTGAGGCTCAACGACAACCCCTGGGTG
TGTGACTGCCGGGCACGCCACTCTGGGCTGGCTGCAGAAGTTCGCGGGCTCCTCTCCGA
GGTGCCCTGCAGCCTCCGCAACGCTGGCTGGCCGTGACCTCAAACGCTAGCTGCCAATG
ACCTGCAGGGCTGCGCTGTGGCCACCGGCCCTTACCATCCCATCTGGACCGGCAGGGCCACC
GATGAGGAGCCGCTGGGGCTTCCCAAGTGCTGCCAGCCAGATGCCGCTGACAAGGCCCTCAGT
ACTGGAGCCTGGAAGACAGCTTCGCGAGGCAATGCGCTGAAGGGACGCTGCCGCCCGGTG
ACAGCCCGCCGGGCAACGGCTCTGGCCACGGCACATCAATGACTACCCCTTTGGGACTCTG
CTTGGCTCTGCTGAGCCCCGCTCACTGCAGTGCGGCCCGAGGGCTCCGAGCCACAGGGTT
CCCCACCTCGGGCCCTCGCCGGAGGCCAGGCTGTTACGCAAGAACCGCACCCGACGCCACT
GCCGTCTGGGCCAGGCAGGCAGCGGGGTGGCGGGACTGGTGACTCAGAAGGCTCAGGTGCC
CTACCCAGCCTCACCTGCAGCCTCACCCCCCTGGGCTGGCGCTGGTGCTGTGGACAGTGCT
TGGGCCCTGCTGACCCCCAGCGGACACAAGAGCGTGCTCAGCAGCCAGGTGTGTGTACATAC
GGGGTCTCTCTCCACGCCGCCAAGCCAGCCGGGCGGCCGACCCGTGGGGCAGGCCAGGCCAG
GTCTCTCCCTGATGGAGCCTGCGGCCCGCCACCCCCCATCTCCACCCCATCATGTTTACAGGG
TTCGGCGGCAGCGTTTGTTCAGAACGCGCCTCCACCCAGATCGCGGTATATAGAGATAT
GCATTTTATTTTACTTGTGTAAAAATATCGGACGACGTGGAATAAAGAGCTCTTTTCTTAAA
AAAA

FIGURE 161

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44184
><subunit 1 of 1, 473 aa, 1 stop
><MW: 50708, pI: 9.28, NX(S/T): 6
MKRASAGGSRL LAWVLWLQAWQVAAPCPGACVCYNEPKVTTSCPQQGLQAVPVGIPAASQRI
FLHGNRISHVPAASFRACRNLTILWLHSNVLARIDAAFTGLALLEQLDLSDNAQLRSVDPA
TFHGLGRLHTLHLDRCGLQELGPGFLFRGLAALQYLYLQDNALQALPDDTFRDLGNLTHLFLH
GNRISSVPERAFRGLHSLDRLLHQNRVAHVHPHAFRDLGRLMTLYLFANNLSALPTEALAP
LRALQYLRINDNPWWCDRCRAPLWAWLQKFRGSSSEVPCSLPQRLAGRD LKRLAANDLQGCA
VATGPYHPIWTGRATDEEPLGLPKCCQPDAAADKASVLEPGRPASAGNALKGRVPPGDSPPGN
GSGPRHINDSPFGTLPGSABEPLTAVRPEGSEPPGFPTSGPRRRPGCSRKNRTRSHCRLGQA
GSGGGGTGDSESGALPSLTCSLTPLGLALVLWTVLGP
```

Important features:

Signal peptide:

amino acids 1-26

Leucine zipper pattern.

amino acids 135-156

Glycosaminoglycan attachment site.

amino acids 436-439

N-glycosylation site.

amino acids 82-85, 179-183, 237-240, 372-375 and 423-426

VWFC domain

amino acids 411-425

FIGURE 162

GGAGTCCACGGGGAGCTTGGATGCCAAGGGAGGACGGCTGGGTCTCTGGAGAGGACTAC
TCAC'TGGCATATTTCTGAGGTATCTGTAGAATAACCACAGCCTCAGATACTGGGGACTTTTAC
AGTCCACAGAACCGTCTCCAGGAAGCTGAATCCAGCAAGAACAATGGAGGCCACGGCGGA
AGCTCAATTTGCGACAAAGGCAAGTCTTTTTCCTTTCTCCTTTGGGCTTATCTCTGGCG
GGCGGGCGGGAACCTAGAGAAGTATTTCTGTGGTGGAGGAAACTGAGGGGCAGCTCCTTTGTGCAC
CAATTTAGCAAAGGACTGGGTCTGGAGCAGAGGGAATTTCCAGCGCGGGGTGTAGGGTTG
TTTCCAGAGGGAACAACTACATTTCAGCTCAATCAGGAGACCGCGGATTTGTTGCTAAAT
GAGAAATTTGGACCGTGAGGATCTGTGCGGTACACAGAGCCTGTGTGCTACGTTTCCAAGT
GTTGCTAGAGAGTCCCTTCGAGTTTTCAGGTTCAGCTGCAAGTGAAGTAAAGACATAAACGACC
ACTCTCCAGTATTTCTGGACAACAAATGTTGGTGAAGTATCAGAGAGCAGTCTCTCTGGG
ACTACGTTTCTCTGAAGAATGCCGAAGACTTAGATGTAGGCCAAAACAATATTGAGAACTA
TATAATCAGCCCCAACTCCTATTTTCGGGTCTCACCCGCAAAACGCAGTGATGGCAGGAAAT
ACCCAGAGCTGGTGCTGGACAAGCGCTGGACCGAGAGGAAGAAGCTGAGCTCAGGTTAAACA
CTCACAGCACTGGATGGTGGCTCTCCGCCAGATCTGGCACTGCTCAGGTCTACATCGAAGT
CCTGGATGTCAACGATAATGCCCCGTAATTTGAGCAGCCTTTCTATAGAGTGCAGATCTCTG
AGGACAGTCCGGTAGGCTTCTCGTGTGTGAAGTCTCTGCCACGGATGTAGACACAGGAGTC
AACGAGAGCAATTTCTATTCACTTTTCCAGCTTCAGAAGAGATTTGGCAAAAACCTTTAAGAT
CAATCCTTTGACAGGAGAAATGAACATAAAAAACAACCTCGATTTTGA AAAA ACTTTCAGTCTCT
ATGAAGTCAATATTGAGGCAAGAGATGCTGGAACCTTTTCTGGAATGCACCGTTCTGATT
CAAGTGATAGATGTGAACGACCATGCCCAAGAAGTTACCATGTCTGCATTTACCGACCAAT
ACCTGAGAACCGCCCTGAAACTGTGGTTGCAC'TTTTCAGTGTTCAGATCTTGATTGAGGAG
AAAAATGGGAAAAATTAGTTGCTCCATTCAGGAGGATCTACCCCTTCTCTGAAATCCGCGGAA
AACTTTTACACCTACTAACGGAGAGACCACTAGACAGAGAAAGCAGAGCGGAATACAACAT
CACTATCACTGTCACTGACTTGGGGACCCCTATGCTGATAACACAGCTCAATATGACCGTGC
TGATCGCCGATGTCAATGACAACGCTCCCGCCTTACCCAAACCTCTACACCTGTTCGTG
CGCGAGAACACAGCCCCGCCCTGCACATCCGCAGCGTCAGCGCTACAGACAGAGACTCAGG
CACCAACGCCAGGTCACTACTCGTCTGCTGCCGCCCCAGGACCGCACCTGCCCCCTACAT
CCCTGGTCTCCATCAACGCGGACAAACGGCCACCTGTTGCCCCCTCAGGTCTCTGGACTACGAG
GCCCTGCAGGGGTTCCAGTTCGCCGTGGGCGCTTCAGACCACGGCTCCCCGGCGCTGAGCAG
CGAGGCGCTGGTGCGGTGGTGGTCTGGACGCCAAGCAACAACCTCGCCCTCTGTTGCTACC
CGCTGCAGAACGGCTCCGCGCCCTGCACCAGAGCTGGTGCCCCGGGCGGCCAGCGGGGCTAC
CTGGTGACCAAGTGGTGGCGGTGGACGGCGACTCGGGCCAGAACGCCCTGGCTGTGCTACCA
GCTGTCTCAAGGCCACGAGCTCGGTCTGTTCCGCGTGTGGGCGCACAAATGGCCAGGTGCGCA
CCGCCAGGCTGCTGAGCGAGCGACGACGGGCCAAGCAAGGCTGGTGGTGTGCTGGTCAAGGAC
AATGCGAGCCTCCGCGCTCGGCCACCGCCACGCTGCACGTGCTCCTGTGGAGCGGCTTCTC
CAGGCCCTACCTGCCCTCTCCCGGAGGCGGCCGACCCAGGCCAGGCCGACTTGTCTCACCG
TCTACCTGGTGTGGCGTTAGGCCCTCGGTGTCTTCGCTCTTCTCTTTTCGGTGCTCTCTGTTC
GTGGCGGTGCGCTGTGTAGGAGGAGCAGGGCGGCCCTCGGTGGTGTGCTGGTGTGCGCCGA
GGGCCCCCTTCCAGGGCATCTTGTGGACATGAGCGGCACCAGGACCCCTATCCAGAGCTACC
AGTATGAGGTGTGTCTGGCAGGAGGCTCAGGGACCAATGAGTTCAAGTTTCTGAAAGCCGATT
ATCCCCAACTTCCCTCCCCAGTGCCCTGGGAAAGAAATACAAGGAAATCTACCTCTCCCAA
TAAC'TTGGGTTCAATATTAGTGAACCATAGTTGACTTTTACATTCCATTAGGTATTTTATTT
TGTGGCATTTCATGCCCATGTTTATTTCCCCCAATTTGTGTGTATGTAATATTGTACGGAT
TTACTCTTGATTTTCTCATGTCTTTCTCCCTTTGTTTAAAGTGAACATTAACCTTTTAT
CCTGGTTCTT

FIGURE 163

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48314
<subunit 1 of 1, 798 aa, 1 stop
<MW: 87552, pI: 4.84, NX(S/T): 5
MEASGKLICRQQLVFSLLLGLSLAGAAEPRSYSVVEETEGSSFVTNLAKDLGLEQREFSR
RGVRVVS RGNKLHLQLNQETADLLLNEKLDREDLCGHTEPCVLRFPQVLLES PFEFFQAE LQV
IDINDHSPVFLDKQMLVKVSESSPPGTTFFPLKNAEDLDVGQNNIENYIISPNSYFRVLTRKR
SDGRKYPELVLDKALDREEEAELRLTLTALDGGSPPRSGTAQVYIEVLVDNDNAPEFEQPFY
RVQISEDSPVGFVLVVKVSATDVDVTGVNGEISYSLFQASEEIGKTFKINPLTGEIELKKQLDF
EKLQSYEVNIEARDAGTFSGKCTVLIQVIDVDNHAPEVTMSAFTSPIPENAPETVVALFSVS
DLDSGENGKISCSIQEDLPFLLKSAENFYTLTTERPLDRESRAEYNITITVTDLGTFMLITQ
LNMVTLIADVDNDNAPAFQTQTSYTLFVRENNSPALHIRSVSATDRDSGTNAQVTYSLLPQDP
HLPLTSLVSINADNGHLFALRSLDYEALQGQFQFRVGASDHGSPALSSSEALVRVVVLDANDNS
PFVLYPLQNGSAPCTELVPRAAEPGYLVTKVVAVDGDSGQNAWLSYQLLKATELG LFGVWAH
NGEVRTARLLSERDAAKHRLVVLVKDNGEPPRSATATLHVLLVDGFSQPYLPPEAAPTQAQ
ADLLTVYLVVALASVSSLFLFVSVLLFVAVRLCRRSRAASVGRCLVPEGFLPGHLVDMSGTRT
LSQSYQYEVCLAGGSGTNEFKFLKPIIPNFPQCPCGEIQGNSTFPNNFGFNIQ

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 685-712

Cadherins extracellular repeated domain signature.

amino acids 122-132, 231-241, 336-346, 439-449 and 549-559

ATP/GTP-binding site motif A (P-loop).

amino acids 285-292

N-glycosylation site.

amino acids 418-421, 436-439, 567-570 and 786-789

FIGURE 164

ACCCACGCGTCCGCCACGCGTCCGCCACGCGTCCGCCACGCGTCCGCGCGTAGCCGTGC
GCCGATTGCCTCTCGGCCTGGGCAATGGTCCCGGCTGCCGGTCGACGACCGCCCCGCGTCAT
GCGGCTCCTCGGTGTTGGCAAGTATTGCTGTGGGTGCTGGGACTTCCCGTCCGCGGCGTGG
AGGTTGCAGAGGAAAGTGGTCGCTTATGGTCAGAGGAGCAGCCTGCTCACCTCTCCAGGTG
GGGGCTGTGTACCTGGGTGAGGAGGAGCTCCTGCATGACCCGATGGGCCAGGACAGGGCAGC
AGAAGAGGCCAATGCGGTGCTGGGGCTGGACACCCAAGGCGATCACATGGTGATGCTGTCTG
TGATTCTCTGGGAAGCTGAGGACAAAGTGAGTTCAGAGCCTAGCGGCGTCACCTGTGGTGCT
GGAGGAGCGGAGGACTCAAGGTGCAACGTCCGAGAGAGCCTTTTCTCTCTGGATGGCGCTGG
AGCACACTTCCTTGACAGAGAAGAGGAGTATTACACAGAGCCAGAAGTGGCGGAATCTGACG
CAGCCCCGACAGAGGACTCCAATAACACTGAAAGTCTGAAATCCCCAAAGGTGAACTGTGAG
GAGAGAAACATTACAGGATTAGAAAAATTTCACTCTGAAAAATTTTAAATATGTCACAGGACCT
TATGGATTTTCTGAACCCAAACGGTAGTGACTGTACTCTAGTCTGTGTTTTACACCCCGTGGT
GCCGCTTTTCTGCCAGTTTGGCCCTCACTTTAACTCTCTGCCCGGGGCATTTCCAGCTCTT
CACTTTTGGCACTGGATGCATCTCAGCACAGCAGCCTTTCTACCAGGTTTGGCACCCTAGC
TGTTCCCTAATATTTATTATTTCAAGGAGCTAAACCAATGGCCAGATTTAATCATACAGATC
GAACACTGGAAACACTGAAAATCTTCATTTTTAATCAGACAGGTATAGAAGCCAGAAGAAT
GTGGTGGTAACTCAAGCCGACCAAATAGGCCCTCTTCCAGCACTTTGATAAAAAGTGTGGA
CTGGTTGCTTGTATTTTCTTATTCTTTTTAATTAGTTTTATTATGTATGCTACCATTCGAA
CTGAGAGTATTCGGTGGCTAATTCAGGACAAGAGCAGGAACATGTGGAGTAGTGATGGTCT
GAAAGAAGTTGGAAAGAGGAACTTCAATCCTTCGTTTCAGAAATTAGTGCTACAGTTTCATA
CATTTTCTCCAGTGACGTGTTGACTTGAAACTTCAGGCAGATTAAGAAGATCATTTGTTGAA
CAACTGAATGTATAAAAAAATTATAAACTGGTGTTTTAACTAGTATTGCAATAAGCAAATGC
AAAAATATTCAATAG

FIGURE 165

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48333
><subunit 1 of 1, 360 aa, 1 stop
><MW: 39885, pI: 4.79, NX(S/T): 7
MVPAAAGRRPPRMRLLGWWQVLLWVLGLPVRGVEVAEESGRLWSEEQPAHPLQVGAVYLGEE
ELLHDPMGQDRAAEEANAVLGLDTQGDHVMVLSVIPGEAEDKVSSEPSGVTTCGAGGAEDSRC
NVRESLFSLDGAGAHFPDREEEYYTEPEVAESDAAPTEDSNNTESLKSPKVNCEERNITGLE
NFTLKILNMSQDLMDFLNPNGSDCTLVLFYTPWCRFSASLAPHFNSLPRAFPALHFLALDAS
QHSSLSTRFGTVAVPNILLFQGAKPMARFNHTDRTLLETCLKIFIFNQTGIEAKKNVVVTQADQ
IGPLPSTLIKSVDWLLVFSLSLFLISFIMYATIRTESIRWLIPGQE QEHE
```

Important features:

Signal peptide:

amino acids 1-25

Transmembrane domain:

amino acids 321-340

Homologous region to dilsufide isomerase

amino acids 212-302

N-glycosylation site.

amino acids 165-168, 181-184, 187-190, 194-197, 206-209, 278-281
and 293-296

Thioredoxin domain

amino acids 211-227

FIGURE 166

CCCGGCTCCGCTCCCTCTGCCCCCTCGGGGTCGCGCGCCACGATGCTGCAGGGCCCTGGCT
CGCTGCTGCTGCTCTTCTCGCCTCGCACTGCTGCCTGGGCTCGGCGCGGGGCTCTTCCTC
TTTGGCCAGCCGACTTCTCTACAAGCGCAGCAATTGCAAGCCCATCCCGGTCAACCTGCA
GCTGTGCCACGGCATCGAATACCAGAACATGCGGCTGCCCAACCTGCTGGGCCACGAGACCA
TGAAGGAGGTGCTGGAGCAGGCCGCGCTTGATCCCGCTGGTCATGAAGCAGTGCCACCCG
GACACCAAGAAGTTCCTGTGCTCGCTCTTCGCCCCGCTGCTCGATGACCTAGACGAGAC
CATCCAGCCATGCCACTCGCTCTGCGTGCAGGTGAAGGACCGCTGCGCCCCGGTCATGTCCG
CCTTCGGCTTCCCTTGCCCCGACATGCTTGAGTGCAGCCGTTTCCCCAGGACAACGACCTT
TGCAATCCCCCTCGCTAGCAGCGACCACTCCTGCCAGCCACCGAGGAAGCTCCAAAGGTATG
TGAAGCCTGCAAAAATAAAAAATGATGATGACAACGACATAATGGAAACGCTTTGTAAAAATG
ATTTTGCACTGAAAAATAAAAGTGAAGGAGATAACCTACATCAACCGAGATACCAAAATCATC
CTGGAGACCAAGAGCAAGACCATTTACAAGCTGAACGGTGTGTCCGAAAGGGACCTGAAGAA
ATCGGTGCTGTGGCTCAAAGACAGCTTGCAGTGCACCTGTGAGGAGATGAACGACATCAACG
CGCCTATCTGGTCATGGGACAGAAACAGGGTGGGGAGCTGGTGATCACCTCGGTGAAGCGG
TGGCAGAAGGGGCAGAGAGATTCAAGCGCATCTCCCGCAGCATCCGCAAGCTGCAGTGCTA
GTCCCGGCATCCTGATGGCTCCGACAGGCTGCTCCAGAGCACGGCTGACCATTTCTGCTCC
GGGATCTCAGCTCCCGTTCCCCAAGCACACTCCTAGCTGCTCCAGTCTCAGCCTGGGCAGCT
TCCCCCTGCCTTTTGCACTTTGCACTCCCAGCATTTTCTGAGTTATAAGGCCACAGGAGTG
GATAGCTGTTTTCACCTAAAGGAAAAGCCACCCGAATCTTGTAAGAAATATTCAAACATAA
AAATCATGAATATTTTAA

FIGURE 167

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50920
><subunit 1 of 1, 295 aa, 1 stop
><MW: 33518, pI: 7.74, NX(S/T): 0
MLQGPGSLLLLFLASHCCLGSARGLFLFGQPDFSYKRSNCKPIPVNLQLCHGIEYQNMRLPN
LLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAPVCLDDLDLDETIQPCHSLCVQVKDR
CAPVMSAFGFPWPDMLECDRFPQDNDLCIPLASSDHLLPATEEAPKVCEACKNKNDNDNDIM
ETLCKNDFALKIKVKEITYINRDTKIILETKSKTIYKLVGVSEKDLKKSVLWLKDSLQCTCE
EMNDINAFYLVMGQKQGELVITSVKRWQKGQREFKRISRIRKLQC
```

Important features:

Signal peptide:

amino acids 1-20

Cysteine rich domain, homologous to frizzled N terminus

amino acids 6-153

FIGURE 168

GTGGAGGCCGCCGACGATGGCGGGGCCGACGAGGCCGAGACGGGGTTGGCCGAGCCCCGGG
CCCTGTGCGCGCAGCGGGGCCACCGCACCTACGCGCGCCGCTGGGTGTTCTGTCTCGCGATC
AGCCTGTCTCAACTGTCTCCAAACGCCACGCTGTGGCTCAGCTTTGCACCTGTGGCTGACGTCA
TGCTGAGGACTTTGGTCCTGTCCATGGAGCAGATCAACTGGCTGTCACTGGTCTACCTCGTGG
TATCCACCCCATTTGGCGTGGCGGCCATCTGGATCCTGGACTCCGTCGGGCTCCGTGCGGCG
ACCATCCTGGGTGCGTGGCTGAACCTTTGCCGGGAGTGTGTACGCATGGTGCCTGTCATGGT
TGTTTGGGACCCAAAACCATTTGCTTCTCTCATGGGTGGCCAGAGCCTCTGTGCCCTTGCCC
AGAGCCTGGTCATCTTCTCTCCAGCCAAGCTGGCTGCCTTGTGGTTCCAGAGCACACGCGA
GCCACGGCCAACATGTCTCGCCACCATGTCTGAACCTCTGGGCGTCTTGTGGCCAATGTGCT
GTCCCTGTGCTGGTCAAGAAGGGTGAGGACATTCGGTTAATGTCTCGTGTCTATACCATCC
CTGCTGGCGTCTGTCTGCCTGCTGTCCACCATCTGCCTGTGGGAGAGTGTGCCCCCACC
CCCTCTGCCGGGGCTGCCAGCTCCACCTCAGAGAAGTTCTCGGATGGGCTCAAGCTGCAGCT
CATGTGGAACAAGGCCTATGTCTATCCTGGCTGTGTCTTGGGGGAATGATCGGGATCTCTG
CCAGCTTCTCAGCCCTCCTGGAGCAGATCCTCTGTGCAAGCGGCCACTCCAGTGGGTTTCC
GGCCTCTGTGGCGCTCTTTCATCACGTTTGGGATCCTGGGGGCACTGGCTCTCGGCCCTTA
TGTGGACCGGACCAAGCACTTCACTGAGGCCACCAAGATTGGCCTGTGCCTGTCTCTCTGG
CCTGCGTGCCCTTTGCCCTGGTGTCCAGCTGCAGGGACAGACCCTTGCCCTGGCTGCCACC
TGCTCGTCTCGGGCTGTTTGGCTTCTCGGTGGGCCCCGTGGCCATGGAGTTGGCGGTGCA
GTGTTCTTCCCCGTGGGGGAGGGGGCTGCCACAGGCATGATCTTTGTGCTGGGGCAGGCCG
AGGGAATACTCATCATGCTGGCAATGACGGCACTGACTGTGCGACGCTCGGAGCCGTCTTG
TCCACCTGCCAGCAGGGGAGGATCCACTTGACTGGACAGTGTCTCTGCTGCTGATGGCCGG
CCTGTGCACCTTCTTCACTGTCATCCTGGCGGTCTTCTTCCACACCCCATACCGGCGCCTGC
AGGCCGAGTCTGGGGAGCCCCCTCCACCCGTAACGCCGTGGGCGGCGCAGACTCAGGGCCG
GGTGTGGACCGAGGGGGAGCAGGAAGGGCTGGGGTCTGGGGCCAGCACGGCGACTCCGGA
GTGACCGGCGAGGGGGGCTCGCTAGAGGACCCAGAGGGCCCGGAGCCCCACCCAGCCT
GCCACCGAGCGACTCCCCGTGCGCAAGGCCAGCAGCCACCGACGCGCCTCCCGCCCCGGC
AGACTCGCAGGCGAGGGTCCAAGCGTCCAGGTTTATTGACCCGGCTGGGTCTCACTCCTCTT
CTCTCCCCGTGGGTGATCACGCTAGCTGAGCGCCTTGTAGTCCAGGTTGCCCGCCACATCGA
TGGAGGCGAATCGAATCTGGTCCACCTGCGGGCGGGGGCGAAAGGGCTCTTGGCGGCT
CCGGGAGCGAATTACAAGCGCGCACCTGAAAA

FIGURE 169

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50988
><subunit 1 of 1, 560 aa, 1 stop
><MW: 58427, pI: 6.86, NX(S/T): 2
MAGPTEAETGLAEPRALCAQRGHRTYARRWVFLLAISLLNCSNATLWLSFAPVADVIAEDLV
LSMEQINWLSLVYLVVSTPPFGVAAIWILDSVGLRAATILGAWLNFAGSVLRMVPCMVVGTQN
PFAPLMGGQSLCALAQSLVIFSPAKLAALWFPEHQRATANMLATMSNPLGVLVANVLSPLV
KKGEDIPMLGVYITPAGVVCLLSTICLWESVPPTPPSAGAASSTSEKFLDGLKLQLMWKNA
YVILAVCLGGMIGISASFALLEQILCASGHSSGFSGLCGALFITFGILGALALGPYVDRTK
HFTEATKIGLCLFSLACVPFALVSQLQGQTLALAATCSLLGLFGFSVGPVAMELAVECSFPV
GEGAAATGMIFVLGQAEIGILIMLMTALTVRREPSLSLTCQQGEDPLDWTVSLLLMAGLCTFF
SCILAVFFHTPYRRLQAESGEPSPSTRNAVGGADSGPGVDRGGAGRAGVLGPSTATPECTARG
ASLEDPRGPGSPHPACHRATPRAQGAATDAPSRPGRLAGRVQASRFIDPAGSHSSFSSPWVIT

Important features:

Signal peptide:

amino acids 1-44

Transmembrane domains:

amino acids 61-79, 98-112, 126-146, 169-182, 201-215, 248-268,
280-300, 318-337, 341-357, 375-387, 420-441

N-glycosylation site.

amino acids 40-43 and 43-46

Glycosaminoglycan attachment site.

amino acids 468-471

FIGURE 171

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48331
<subunit 1 of 1, 1184 aa, 1 stop
<MW: 129022, pI: 5.20, NX(S/T): 5
MMQLQLLLGLLGGPGGYLFLLGDCQEVTTLTVKYQVSEEVPSGTVIGKLSQELGREERRRQA
GAAPQVLQLPQALPIQVDSEGLLSTGRRLDREQLCRQWDPCLVSFDVLATGDLALIHVEIQ
VLDINDHQPRFPKGEQELEISESASLRTRIPLDRALDPDTGPNLTHTYTLSPSEHFALDVIV
GPDETKHAEILIVVKELDREIHSFFDLVLTAYDNGNPPKSGTSLVKVNVLDSDNSPAPAESS
LALEIQEDAAPGTTLLIKLTATDPDQGPNGEVEFFLSKHMPPVELDTFSIDAKTGQVILRRPL
DYEKNPAYEVDVQARDLGPNIPIAHCKVLIKVLVDVNDNIPSIHVTWASQPSLVSEALPKDSF
IALVMADDLDSGHNGLVHCWLSQELGHFRLKRTNGNTYMLLTNATLDREQWPKYTLTLLAQD
QGLQPLSAKKQLSIQISDINDNAPVFEKSRYEVSTRENNLPSLHLITIKAHADLGINGKVS
YRIQDSPVAHLVAIDSNTGEVTAQRSNLNYEEMAGFEFQVIAEDSGQPMLASSVSVVWSLLDA
NDNAPEVVQPVLSDGKASLSVLVNASTGHLLVPIETPNGLGPAGTDTPTPLATHSSRPFLTTT
IVARDADSGANGEPLYSIRNGNEAHLFILNPHTGQLFVNVTINASSLIGSEWELEIVVEDQGS
PPLQTRALLRVMFVTSVDHLRDSARKPGALSMMLTVICLAVLLGIFGLILALFMSICRTEK
KDNRAYNCREAESTYRQPKRPQKHIIKADIHLVPVLRGQAGEPCEVGQSHKDVCKEAMMEA
GWDPCQLQAPFHILTPTLYRTLNRQGNQGAPAESREVLQDVTNLLFNHPRQRNASRENLLPEP
QPATGQPRSRPLKVAGSPTGRLAGDQGSSEAPQRPASSATLRRQRHLNGKVSPEKESGPRQ
ILRSLVRLSVAFAERNPVEELTVDSPPVQQISQLLSLLHQGFQPKPNHRGNKYLAKPGGS
RSAIPDTDGPSARAGGQTDPEQEGLDPEEDLSVKQLLEEELSSLLDPSTGLALDRLSAPD
PAWMARLSLPLTTNYRDNVISPDAAATEEPRTFTQFGKAEAPELSPTGTRLASTFVSEMSSL
LEMLLEQRSSMPVEAASEALRRLSVCGRTLSDLATSAASGMKVQGDPGGKTGTGKSRGSS
SSSRCL
```

Important features:

Signal peptide:

amino acids 1-13

Transmembrane domain:

amino acids 719-739

N-glycosylation site.

amino acids 415-418, 582-585, 659-662, 662-665 and 857-860

Cadherins extracellular repeated domain signature.

amino acids 123-133, 232-242, 340-350, 448-458 and 553-563

FIGURE 172

CGGACGCGTGGGCGGACGCGTGGGGGAGAGCCGCGAGTCCCGGCTGCAGCACCTGGGAGAAGG
CAGACCGTGTGAGGGGGCCTGTGGCCCCAGCGTGCTGTGGCCTCGGGGAGTGGGAAGTGGAG
GCAGGAGCCCTTCCTTACACTTCGCCATGAGTTTTCCTCATCGACTCCAGCATCATGATTACCT
CCCAGATACTATTTTTTGGATTGGTGGCTTTTCTTCATGCGCCAATTGTTTAAAGACTAT
GAGATACGTCAGTATGTTGTACAGGTGATCTTCTCCGTGACGTTTGCATTTTCTGCACCAT
GTTTGAGCTCATCATCTTTGAAATCTTAGGAGTATTGAATAGCAGCTCCCGTTATTTTCACT
GGAAATGAACCTGTGTGTAATTCTGCTGATCCTGGTTTTCATGGTGCCTTTTACATGCG
TATTTTATTGTGAGCAATATCCGACTACTGCATAAAACAAGACTGCTTTTTCTGTCTCTT
ATGGCTGACCTTTATGTATTTCTTCTGGAACCTAGGAGATCCCTTTCCCATCTCAGCCCAA
AACATGGGATCTTATCCATAGAACAGCTCATCAGCCGGGTTGGTGTGATTGGAGTGACTCTC
ATGGCTCTTTCTGGATTGGTGTCTCAACTGCCCATACACTTACATGTCTTACTTCTCT
CAGGAATGTGACTGACACGGATATTCTAGCCCTGGAACGGCGACTGCTGCAAAACCATGGATA
TGATCATAAGCAAAAAGAAAGGATGGCAATGGCACGGAGAACAATGTTCCAGAAGGGGGAA
GTGCATAACAAACCATCAGGTTTCTGGGGAATGATAAAAGTGTTACCACTCAGCATCAGG
AAGTGAAATCTTACTCTTATTCAACAGGAAGTGGATGCTTTGGAAGAATTAAAGCAGGCAGC
TTTTTCTGGAACAGCTGATCTATGCTACCAAGGAGAGATAAGATACTCCAAACCTTC
AAGGGGAAATATTTTAAATTTTCTGGTTACTTTTTCTCATTTACTGTGTGTTGGAAAAATTT
CATGGCTACCATCAATATGTTTTTGTATCGAGTTGGGAAAAACGGATCCTGTCAAGAGGCA
TTGAGATCACTGTGAATTATCTGGGAATCCAATTGTAGTGAAGTTTGGTCCCAACACATT
TCCTTCATTCTGTTGGAATAATCATCGTCACATCCATCAGAGGATGCTGATCACTCTTAC
CAAGTTCTTTTATGCCATCTCTAGCAGTAAGTCTCCAATGTCTTGTCTGCTATTAGCAC
AGATAATGGGCATGTACTTTGTCTCCTCTGTGCTGCTGATCCGAATGAGTATGCCTTTAGAA
TACCGCACCATAACTACTGAAGTCTTGGAGAACTGCAGTTCAACTTCTATCACCCTTGGTT
TGATGTGATCTTCTGGTGCAGCGCTCTCTAGCATACTCTTCTCTATTGGGCTCAGAAAC
AGGCACCAGAGAAGCAATGGCACCTTGAACTTAAGCTACTACAGACTGTTAGAGGCCAGT
GGTTTCAAAATTTAGATATAAGAGGGGGGAAAAATGGAACCGGGGCTGACATTTTATAAAC
AAACAAAATGCTATGGTAGCATTTTTTACCTTCATAGCATACTCTTCCCGTCAGGTGATA
CTATGACCATGAGTAGCATCAGCCAGAACATGAGAGGGAGAACTAACTCAAGACAATACTCA
GCAGAGAGCATCCCGTGTGGATATGAGGCTGGTGTAGAGGCGGAGAGGCCAAGAACTAA
AGGTGAAAAATACACTGGAACTCTGGGGCAAGACATGTCTATGGTAGCTGAGCCAAACAGT
AGGATTTCCGTTTTAAGGTTCACTGGAAAAGGTTATAGCTTTGCCTTGAGATTGACTCATT
AAAAATCAGAGACTGTAACAAAAAAAAAAAAAAAAAAAGGGCGGCGGAGCTCTAGAGTCG
ACCTGCAGGAAGCTTGGCCGCCATGGCCCACTTGTATTTCAGCTTTATAATG

FIGURE 173

MSFLIDSSIMITSQILFFGFGWLFFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEI
LGVLNSSSRYPFHWMNLCVILLILVFMVPPYIGYFIVSNIRLLHKQRLLFSCLLWLTFFMYFF
WKLGDPPILSPKHGILSIEQLISRVGIVGVTLMALLSGFGAVNCPYTYMSYFLRNVTDTDI
LALERRLLQTMDMIISKKKRMAMARRTMFQKGEVHNKPSGFWGMIKSVTTSASGSENLTLLIQ
QEVDALBBLSRQLFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVF
DRVGKTDVPVTRGIEITVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTFFFYAISS
SKSSNVIVLLLAQIMGMYFVSSVLLIRMSMPLEYRTIITEVLGELQFNFYHRWFDVIFLVSA
LSSILFLYLAHKQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 174

CATGGGAAGTGGAGCCGGAGCCTTCCTTACACTCGCCATGAGTTTCCTCATCGACTCCAGCA
TCATGATTACCTCCCNGANACTATTTTTTGGATTGGGTGGCTTTTCTTCNGCGCCAATGTT
TAAAGACTATGAGATACGTCAGTATGTTGTACNGGTGATCTTCTCCGTGACGTTTGCCATTT
CTTGCACCATGTTTGAGCTCATCATCTTTGAAATCTTNGGAGTATTGAATAGCAGCTCCCGT
TATTTTCACTGGAAAATGAACCTGTGTGTAATTCTGCTGATCCTGGTTNTCATGGTGCCTTT
TTACATTGGCTATTTTATTGTGAGCAATATCCGACTACTGCATAAACAACGACTGCTTTTTT
CCTGTCTCTTATGGCTGACCTTTATGTATTTCCAG

100
90
80
70
60
50
40
30
20
10
0
100
90
80
70
60
50
40
30
20
10
0

Figure 1

GTGTGTGCCCTTGGGGAGGGGAAGGGGAGCCNGGCCCTTTCTTAAAAATTGGCCAAGGGTTTC
TTTNTTGAATTCGGGTNNNGNATACCTTCCCAGAAAATATTTTTTGGATTGGGGTAGNTT
TTTTTCATGCGCCAAATGTTTAAAGACTATGAGATACGTGATATGTTGTACAGGTGATNTT
NTCCGTGACGTTTGCATTTTCTTGACCATGTTTGAGCTCATCATNTTTGAAATNTTAGGAG
TATTGAATAGCAGCTCCCGTTATTTTCACTGGAATAAGACCTGTGTGTAATTCTGCTGATC
CTGGTTTTTCATGGTGCCTTTTTCTCATTTGGCTATTTTATTGTGAGCAATATCCGACTACTGCA
TAAACAACGACTCTTTTTCCTGTCNTTATGGCTGACCTTTATGTATTTNTTNTGGAAAN
TAGGAGATCCCTTTCCCAATTCCTC

FIGURE 176

CTCGCGCAGGSGATCGTCCC**ATG**CCCGGGCTCGGAGCCGCGACCCTTGGGGGGCCCTCCGGGATTTGCTACCTTTT
 TGGCTCCTCGTCTCGTGAACATGCTCTTCTACAGGGCTGTGCGCTTCAATCTGGAGCGTGATGGGTGCTTGCGCCAA
 GAGGGGCGAGCCAGCGAGCTCTTCGGCTTCTCTGTCGGCCCTGCACCCGGCATTTGCGAGCCCGACGCCCAAGAGCTG
 GCTGCTGTGGGTGCTCCCCAGGCGCTTGGCTTCTCTGCGGAGCAGCGCGAATCGCACTGGAGGCTCTTTCGCTTG
 CCGCTTGAGCTCTGGAGGAGTGAAGTCTGCTACAGAGTGGACATCGACAGGGAGCTGATATGCAAAAGGAAAGCAA
 GAGGAACCAAGTGGTGGGAGTCAAGTGTTCGGAGCCAGGGGCTGGGGCGAAGATTGTACTCTGTGCACACCGATA
 TGAGGCAAGGACGCGAGTGGACCAAGTCTGGAGACCGGGATATGATTGTCTGCTCTTTGCTGCTCAGCCAGGA
 CTCTGGCCATCCGGGATGAGTTGGATGGTGGGGAAATGGAAGTTCTGTGAGGGACGCGCCCAAGGCCATGAACAAAT
 TGGGTTCTGCCAGCAGGGCAGAGCTGCCGCTTCTCCCTGATAGCCACTACCTCTCTTTGGGGCCCTCAGGAAC
 CTATAATTGGAAGGGGACCGCCAGGGTGGAGCTCTGTGCACAGGGCTCAGCGGACCTGGCACAGCTTGGACGAGCGG
 TCCCTACGAGGGCGGGGAGAGAAGGACGAGGACCCCGGCTCATCCGGTCTCTGTCACAGACTTGTGCTTTGGCTT
 CTCTATTGACTCGGGAAAGGTCTGGTGGTGCAGAAAGAGCTGAGCTTTGTGGCTGGAGCCCCCGCGCAACCA
 CAAGGGTGTCTGTGCTCATCTCTGCGCAAGGACAGCGCCAGTGCCTGGTCCCGAGGTTATGCTGTCTGGGGAGCG
 CCTGACCTCCGGCTTTGGCTACTCACTGGCTGTGGCTGACCTCAACAGTGATGGCTGGCCAGACCTGATAGTGGG
 TGCCCCCTACTCTTTGAGCGCAAGAAGAGCTGGGGGGTCTGTGTATGTACTTGAACCAAGGGGGTCACTG
 GGTGGGATCTCCCTCTCCGGCTCTCGGCTCCCTGACTCATTGTTCCGGATCAGCCTGGCTCTCTCTGGGGA
 CCTCAACCAAGATGGCTTTCCAGATATTGCAAGTGGGTGCCCTTTGATGGTATGGGAAAGTCTCTCATCTACCA
 TGGGAGCAGCCTGGGGTGTCTGCCAAACCTTCAACAGTGTCTGGAGGCGAGGCTGTGGGCTCAAGAGCTTCGG
 CTACTCTCTGTGAGCAGCTTGGATATGGATGGATGGGAACCAATACCTGACCTGTGTGGGCTCTCTGGTGCAGC
 CGAGTGTCTTTCAGGGCCAGACCCATCTCCATGTCTCCATGAGGCTCTTGTTCAGGAAAGTCTGAGAGATCGACCT
 GGACCGCCCAACTCTGTCGGCGCCACTCGGTCTGTGTGGAGCTAAGGGTCTGTTTCAGCTATGTTGAGTCTCC
 CAGCAGCTATAGCCCTATCTGTGGCTGAGCTATGTGTTAGATGCGGACACAGACCGGAGGCTCCGGGGCAGGT
 TCCCCTGTGACGTTCTGTGAGCGTAACTTGGAAAGAACCCAGACACAGGCTCTCGGACCGTGTGGCTGAAGCA
 CCAGCATGACCGAGTCTGTGGAGAGCCATGTTCCAGCTCCAGGAAGATGCTCAAGACAGAGCTTCGGGCAATGT
 AGTGACCTTTGCTTACAGTCTCCAGACCCCTCGGCTCCGGGACAGGCTCTTGCCAGGGGCTGCTCTCAAGTGGC
 CCCCATCCTCAATGCCACCAGCCAGCACCCAGCGGGCAGAGATCCACTTCTGGAAGCAAGGCTGTGGTGAAGA
 CAAGATCTGCCAGAGCAATCTGCAGCTGGTCCAGCCCGCTTCTGTACCCGGGTCAAGCACAGGAAATCCAAAC
 TCTGCCATGGATGGATGGATGGAAACAAAGCCCTGTTTGCACTGAGTGGGCGAGCTCATTTGGCTGGAGCTGAT
 GGTCAACACCTGCCATCGAGCCAGGCCAGGCCAGGCTGATGGGGATGATGCCATGAAGCCAGCTCCTGGT
 CATGCTTCTGACTCACTGCACTACTCAGGGGTCCGGGCTCGGACCTGCGGAGAGGCCACTCTGCTGTCCAA
 TGAGATGCTCCCATGTTGAGTGTGAGCTGGGGAACCCATGAAGAGAGGTGCCAGGTCACTTCTACCTCAT
 CTTTAGCACTCCCGGATCAGCATTGAGACCAGGAACCTGGAGGTAGAGCTGCTGTTGGCCACGATCAGTGAGCA
 GGAGCTGCATCCAGTCTCTGTCAGAGCCCGTGTCTTCAATTGAGCTGCCACTGTCTATTGCGAGGAATGGCCATTCC
 CCAGCAACTCTTCTTCTCTGCTGTGGTGGAGGGGCGAGAGAGCCATCAGCTGAGCGGATGTGGGCGAGCAAGCT
 CAAGTATGAGGTCACGGTTTCAACCAAGGCCAGTGCCTCAGAACCTGGGCTTCGCTTCTCAACATCATGTG
 GCTCATGAGATTGCCAATGGGAAGTGGTGTCTGTACCCAACTCAGGTTTGGAGTGGAGGGGCGGCGAGGCGTGG
 GCGAAGAGGGCTTTGCTCTCCAGGCCCAACATCTCCACTGGATGTGAGCAGTAGGAGTAGGAGGCGGGGGA
 TGTGGAGCTGCTGAGCAGCAGGAGCTGGTGAAGCGCAGGAGCCAGCATGCTCTGTGGCCAGTGTCTCTGTC
 TGAGAAGAAGAAACATCACCTGTGACTGCGCCCGGGGACAGGCCACTGTGTGGTGTTCAGTGCCTCACTCA
 CAGCTTTGACCGCGCGCTGTGCTGCATGTCTGGGCGCTCTTGGAACAGCACTTCTGAGGAGATGATCAGT
 GTGGAAGTCCCTGGAAAGTATTGTCCGGGCCAACATCAGTAGAAGTCTCCATTAAGAAATTTGATGCTCCAGA
 TGCCTCCACAGTGATCCAGTAGGAGTATCTTGGACCCATGGCTGTGGTGGCAGAGGATGCTCTGGAGGGTGGT
 CATCTCTGCTGTACTGGCTGGGCTGTCTGTGTAGCACTGCTGTGCTGCTCTCTGTGGAAGATGGGATTCCTT
 CAACGGGCGAAGCCCGGAGGCCACCTGCCCCAGTACCATTGCGGTGAAGTATGCTCTGGGAAGACCGACAGCA
 GTTCAAGGAGGAGAAGACGGGCACATCTCTGAGGAACAACTGGGGCAGCCCCGGCGGAGGGCCGAGTGACCA
 CCCCATCTGACTGCTGAGCGGCATCCGAGCTCGGCCCTGAGTGGCATCCAGGCGCAGGCAACCGC**TAG**GTTC
 CATGTCCAGCACTGGCTGTGGCTGCCATCCCTTCCCAAGAGATGGCTCTTGGGATGAAGAGGGTATGAGT
 GCGCTGCTGGTGTGCATCAAGATTGGCAGGATCGGCTTCTCAGGGGCAAGAGCTCTCCACCCCAAGAAC
 TCTCTCCACCAACTTCCCTCTAGAGTCTGTGAGATGAGAGTGGTAAATCAGGACAGGGCTGGGGTAGGG
 TGAGAAGGGCAGGGGTGCTGTGTCGAAGGTGGGGAGAAGGGATCTTAATCCCTTCTCTCCATTCAACCTGT
 GTAAAGGACCCAGAGGAGCTGCTCCCGGAAGTGTCTTAACCTAGAGGGTGGGGAGGAGGTGTGTCACTGA
 CTCAGGCTGCTCTCTCTGATTTCGCCCTCTCATCTGACCTTAGTTTGTCTGCCATCAGTCTAGTGGTTTCTGGT
 TTCTGCTATTATTAAAAAATATTGAGAACAAAAA

FIGURE 177

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA55737

><subunit 1 of 1, 1141 aa, 1 stop

><MW: 124671, pI: 5.82, NX(S/T): 5

MAGARSRDPWGASGICYLFGSLLVELLFSRAVAFNLDVMGALRKEGEPGSLFGFSVALHRQL
QPRPQSWLLVGAPQALALPGQQANRTGGFLFACPLSLEETDCYRVDDIDQGADMQKESKENQWL
GVSVRSGPGGKIVTCAHRYEARQRVDQILETRDMIGRCFVLSQLDAIRDELDDGGGEWKFCGE
RPQGEHQFGFCQQGTAAAFSPDSHYLLFGAPGTYNWKGATARVELCAQGSADLAHLDDGPGYEA
GGEKEQDPRLLIPVANSYFGFSIDSGKGLVRAEELS FVAGAPRANHKGAVVILRKDSASRLV
PEVMLSGERLTSGFGYSLAVADLNSDGWPDLLVIGAPYFFERQEEELGGAVYVYLNQGGHWAGI
SPLRLCGSPDSMFGISLAVLGDNLQDGFDPDIAGAPFDGDGKVFIYHGSLSGVVAKPSQVLE
GEAVGIKSGFYSLSGSLDMDGNQYPDLLVGS LADTAVLFRARFILHVSHEVSIAPRSIDLEQ
PNCAGHGSVCVDLRVCFYSIAVPSSYSPTVALDYVLDA DTRRLRGQVPRVTFLSRNLEEPK
HQASGTVWLKHQHDRVCGDAMFQLQENVKDKLRAIVVTLSSYSLQTPRLRRQAPQGGLPPVAP
ILNAHQPSQRAEIHFLKQCGEDKICQSNLQLVHARFCTRVS DTEFQPLPMDVDGTTALFA
LSGQPVI GLELMVTNLPSDPAQFPQADGDDAHEAQLLVMLPDSLHYS GVRALDPAEKPLCLSN
ENASHVECELGPNPMKGAQVTFYFILSTSGIS IETTELEVELLATISBQELHFVSARARVF
IELPLS IAGMAIPQQLFFSGVVRGERAMQSERDVGSKVKEYEVTVSNQGGSLRTLGS AFLNIM
WPHEIANGWLLYPMQVELEGGQGPQGKGLCS PRPNILHLDVDSDRRRRELEPPEQQEPEGE
RQEPSMSWNPVSSAEKKKNITLDCARGTANCVVFSCLYSDRAAVLHVWGRLWNSTFLEEY
SAVKSL EIVVRANITVKSSIKNMLRLDASTVIVPMVYLDPMMAVVAEGVPWWVILLAVLAGLL
VLALVLVLLWKMGGFFKRAKHPEATVPQYHAVKIPREDRQQFKEEKTGTILRNWNWS PRREGP
DAHPILAADGHPGLGPDGHPGPGTA

Important features:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 1040-1062

N-glycosylation sites.

amino acids 86-89, 746-749, 949-952, 985-988 and 1005-1008

Integrins alpha chain proteins.

amino acids 1064-1071, 384-408, 1041-1071, 317-346, 443-465, 385-407, 215-224, 634-647, 85-99, 322-346, 470-479, 442-466, 379-408 and 1031-1047

FIGURE 178

CGCGCCGGGCGCAGGGAGCTGAGTGGACGGCTCGAGACGGCGCGCGTGCAGCAGCTCCAGA
AAGCAGCGAGTTGGCAGAGCAGGGCTGCATTTCAGCAGGAGCTGCGAGCACAGTGTGGCT
CACAAACAAGATGCTCAAGGTGTGAGCCGTACTGTGTGTGTGTGCAGCCGCTTGGTGCAGTCA
GTCTCTCGCAGCTGCCGCGCGGTGGCTGCAGCCGGGGGGCGGTCCGACGGCGGTAATTTTC
TGGATGATAAAACAATGGCTCACCACAATCTCTCAGTATGACAAGGAAGTCGGACAGTGGAAC
AAATTCGAGACGAAGTAGAGGATGATTATTTCCGCACTTGGAGTCCAGGAAAACCCCTCGA
TCAGGCTTTAGATCCAGCTAAGGATCCATGCTTAAAGATGAAATGTAGTCGCCATAAAGTAT
GCATTGCTCAAGATTCTCAGACTGCAGTCTGCATTAGTCACCGAGGCTTACACACAGGATG
AAAGAAGCAGGAGTAGACCATAGGCAGTGGAGGGGTCCCATATTATCCACCTGCAAGCAGTG
CCCAGTGGTCTATCCAGCCCTGTTTGTGGTTTCAGATGGTCATACCTACTCTTTTTCAGTGCA
AACTAGAATATCAGGCATGTGTCTTAGGAAAACAGATCTCAGTCAAATGTGAAGGACATTGC
CCATGTCCTTCAGATAAGCCCACCGTACAAGCAGAAATGTTAAGAGAGCATGCAGTGACCT
GGAGTTTCAGGGAAGTGGCAAAACAGATTGCGGGACTGGTTCAAGGCCCTTCATGAAAGTGGAA
GTCAAAAACAAGAAGACAAAAACATTGCTGAGGCCCTGAGAGAAGCAGATTTCGATACCAGCATC
TTGCCAATTTGCAAGGACTCACTTGGCTGGATGTTTAAACAGACTTGATACAACTATGACCT
GCTATTGGACCACTGACAGCTCAGAAGCATTACCTTGATAAGAATGAACAGTGTACCAAGG
CATTCTTCAATTCTTGTGACACATACAAGGACAGTTTAAATCTAATAATGAGTGGTGCTAC
TGCTTCCAGAGACAGCAAGACCCACCTTGCCAGACTGAGCTCAGCAATATTTCAGAAGCGGCA
AGGGGTAAAGAAGCTCCTAGGACAGTATATCCCCCTGTGTGATGAAGATGGTTACTACAAGC
CAACACAATGTCATGGCAGTGTGGACAGTGTGGTGTGTTGACAGATATGGAATGAAGTC
ATGGGATCCAGAATAAATGGTGTGTCAGATTGTGCTATAGATTTTGAGATCTCCGGAGATTT
TGCTAGTGGCGATTTTCATGAATGGACTGATGATGAGGATGATGAAGACGATATTATGAATG
ATGAAGATGAAATGAAGATGATGATGAAGATGAAGGGGATGATGATGATGATGGTGGTATGAC
CATGATGTATACATTGATGATGACAGTTGAAATCAATAAATTTCTACATTTCTAATATTTA
CAAAAATGATAGCCTATTTAAATATCTTCTCCCCAATAACAAAATGATTCTAAACCTCA
CATATATTTGTATAATTATTTGAAAAATTGCAGCTAAAGTTATAGAACCTTATGTTTAAAT
AAGAATCATTGCTTTGAGTTTTATATTCCCTTACACAAAAAGAAAATACATATGCAGTCTA
GTCAGACAAAATAAAGTTTTGAAGTGCTACTATAATAAATTTTTCACGAGAACAAACTTTGT
AAATCTTCCATAAGCAAAATGACAGCTAGTGCTTGGGATCGTACATGTTAATTTTTTGAAAG
ATAATTCTAAGTGAAATTTAAATAAATAAATTTTTAATGACCTGGGTCTTAAGGATTTAGG
AAAAATATGCATGCTTTAATTGCATTTCCAAAGTAGCATCTTGCTAGACCTAGATGAGTCAG
GATAACAGAGAGATACCATGACTCCAAAAA

FIGURE 179

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49829
><subunit 1 of 1, 436 aa, 1 stop
><MW: 49429, pI: 4.80, NX(S/T): 0
MLKVSAVLCVCAAAWCSQSLAAAAVAAGGRSDGGNFLDDKQWLTTISQYDKEVGQWNKFR
DEVEDDYFRTWSPGKPFQALDPAKDPCLMKCSRHKVCIAQDSQTAVCISHRRLTHRMKEA
GVDHRQWRGPILSTCKQCPVVYPSPVCGSDGHTYSFQCKLEYQACVLGKQISVKCEGHCPCP
SDKPTSTSRNVKRACSDLEFREVANRLRDWFKALHESGSQNKKTTLRLRPERSRFDTSILPI
CKDSLGMFMFNRLDTNYDLLLDQSELRSIYLDKNEQCTKAFFNSCDTYKDSLISNNEWCYCFQ
RQODPPCQTELSNIQKRQGVKKLLGQYIPLCDEEDGYKPTQCHGSVGQCWCVDRYGNEVMGS
RINGVADCAIDFEISGDFASGDFHEWTDDEDEDDIMNDEDEIEDDEDEGDDDDGGDDHDVYI
```

Important features:

Signal peptide:

amino acids 1-16

Leucine zipper pattern.

amino acids 246-267

N-myristoylation sites.

amino acids 357-362, 371-376 and 376-381

Thyroglobulin type-1 repeat proteins

amino acids 353-365 and 339-352

FIGURE 180

CAGACTCCAGATTTCCTCTGTCAACCAGGAGTCCAGAGAGGAAACGCGGAGCGGAGACAACAGTACCTGCACGC
 CTCCTTCAGCCCGGAGTCGCCCCCAGCAGGGATGCGGCGACAAAGATCTGGCTGCCCTCCCCGCTCCTCTCGGCC
 GCTCTGCTCCGCGTCTGCTGCTGCTCGGCCGCGCGCGCTTCAACCTTCCTCGATAGCGACTTCACCTTTACCCCT
 CCCGCGCGCCAGAGGAGTGCTTCTACAGCCCATGCCCTGAAGGCCCTCGCTGGAGATCGAGTACCAAGTTTAA
 GATGGAGCAGGATTAGATATTTGATTTCCATCTTGCTCTCCAGAAGGCAAAACCTTAGTTTGTGAACAAAGAAATA
 TCAGATGGAGTTTCACATCTGTAGAGACTGAAGTTGGTGATACATGTTCTGCTTTGACAATACATTCAGCACCATTT
 TCTGAGAAGGTGATTTTCTTGAATTAATCCTGGATAATATGGGAGAACAGGCAAGAAACAAGAGATTGGAAG
 AAATATATTACTGGCACAGATATATTTGGATATGAACTGGAAGACATCCTGGAATCCATCAACAGCATCAAGTCC
 AGACTAGCAAAAGTGGGCACATACAAATTCCTGCTTAGAGCATTTGAAGCTCTGATCGAAACATACAGAAAGC
 AACTTTGTAGAGACTCAATTTCTGGTCTATGGTTAATTTAGTGGTCTAGTGGTGCTCGACCATTCAGTTTAT
 ATGCTGAAGAGTCTGTTTGAAGATAGAGGAAAAGTAACTTAACCTCAAACTAGAGTCAAGTAACTTGA
 AATGAGGCATAAAATGCAATAAATCTGTACAGTCAAGACCATTAATGGCTCTTCCAAAAATTTTGTAGATATA
 AAAGTAGGAAACAGGTATAATTTTAATCTGAAATTAAGTCTTCACTTTCTGTGCAAGTAACTCTGCTGATCCAG
 TTGTACTTAAGTGTGTAAACAGGAATATTTTGCAGAAATATAGGTTTAACTGAATGAAGCCATATTAATACTGCAT
 TTTCTTAACTTTGAAAAATTTTGCAAATGCTTAGGTGATTTAAATAAATGAGTATTTGGGCCCTAATGTCAACACC
 AGTCTGTTTTTAAACAGGTTCTATTACCCAGAACTTTTTGTAAATGCGGCGAGTTACAAATTAACGTGGGAAGTTT
 TCAGTTTTTAAGTTATAAATCACTGAGAATTACCTTAATGATGGATTGAATAAATCTTTAGAGTACAAAAAGCCCAA
 CTTTTCTCTATTACATATGACTCTCTCTATAATGTAAATAGAATAATAGCTTTGAAATACAATTAGGTTTTTGT
 AGTTTTTATAACCAATACATTTTCAGTGTAAACATATTAGCAGAAAGCATTAGTCTTTGTACTTTGCTTACATTT
 CCAAAGCTGTACATTTTCAGACTCTTAAAAACACAAAGTTTACATTAATAAATTAGGACATGTTTTCTCTTTT
 AAATGAAGAAATATAGTTTAAAAGCTTCTCTCCATAGGGACAATTTTCTCAACCCCTTAACTAAAGTGTAGGA
 TTTTAAAAATTAAGTGTGAGGTAAAAATAGTTTATTTTAAATAGTATCTGTCAAGTTAATATCTGTCAACAGTTAA
 TAATCATGTTATGTTAATTTTAACTGATTTGCTGACTTGGATAATTCATTTATACAGCAGTTTGAAGGAAATA
 TTGCTAAATGATCTGGGCCATCCATAAATAATATCTCCTTTTCTGAGCTCTAAGAAATATCAGAAAAACAGGAA
 AGAATTTAGAAAACTTGAGAAAACTTAACTCAAAATAAAATTCACCTTAAGTAGAACTATAAATAAATATCTAGA
 ATCTGACTGGCTCATCATGACATCTACTCATAACATAAATCAAAGGAGATGATTAATTTCCAGTTAGCTGGAAG
 AAATCTGGCTGTAGGTTTTTATTTTCTACAGAAATCTGGTTTGAATTTATTTTGAAGCAGGTACATTTTATA
 AAATGTAAGCCCTACTGTAAGGTTTAGCACTGGGTGTACATATTTATTAATAAATTTTATTAACAACTTTTAT
 TAAATAGGCCCTTCTGAAACACTTTATTTATGATGTTGAAGTAAGGATTAGAAAACATAGACTCCCAAGTTTAAAA
 CACCTAAATGTGAATAACCCATATATACAACAAAGTTTCTGCCATCTAGCTTTTGAAGTCTATGGGGGCTTAC
 TCAAGTACTAGTAATTTAACTTCATCATGAATGAACATAAATTTTAAAGTTATGCCATTTTATACGGTGTGTTAT
 GACTACATTTGTAGTTAGAAAACAACTTAAATTTGGGTTATAGAACCCCTCAACAGGTTAGTAAATGCTGGAAAT
 CTTGATGAGCAATATGATAACAGAGAGTGATTTTCACTTACACTCATAGTAGTATAAAGAGGATACATTTTCCC
 TCTTAGGCCCTTGGGAGAAGAGCAGCTTAGATTTTCCCTACTGGCAAGGTTTAAAAATGAGGTAAAGCCGAT
 ATGATCAATTAACCTTAATTTGGCCAAGAAAATGCTTCAGGTGCTAGGGGTTATCTCTCGCAACACTTGCAGACAA
 AGGCTCAATAAGCTTCTGCTCATTAAGTACCCCTCCCTTTTGGCGCTGTAAATTTGCAATGAGAAGCAAAATTAACA
 GTACCATAACTAAATGAGCAGGGTACAGATATAAACTACTGCACTCTTTCTATAAACTGTGATTAAGAAATCTA
 CCTCTCTGTATGGCTGTACTGTACTGTACTGTACTGTACTGTACTGTACTGTACTGTACTGTACTGTACTGTACTGT
 ACATGTATGATTTTGTGCCATGATCTTAAACCTATGATTCAGTAACTCTTACCATAAATAAGCAGATATTTGAAGG
 TATTTGGAAAAAGATTTAGGAATACTAAGGACAATTTATTTTATAGACAAAGTAAAAAGCAGATATTTGAAGG
 CATAACAAAAAGCAAAACTTTGTAACAGAGTAAAAATCTTAAATTTCTAAAGACACTCTGTTTATCTGCT
 CATACTCTTTTTTAAATTTTCACTATTCATTCTTAAATTTAAAGTTATGCTAAATTTGAGTAAGCTGTTTATCACTT
 AAAGCATCATTTTGTCTTTTCAATATACAAATTTAAAAAATACTACAAATTTAACTAAGGCCAACCGGATTTTCA
 CATAAGGTAGCAGTACCCTGTTCACCTCACACTAAGGCCATAGATTTGCTGATATGCAATTTGGATGATTAAT
 GTTATGCTGTTCTTTTCACTGTAAGTCAAGACATGGAGGTTGTTGTAATTTATGGTGAATAATTAATCTCTTCA
 CACATAAGTGTGCTTAAATTTGACAAAAAATGAGCACTTACAATTTGATGCTCTTCAATGAAGGATTTCTTAT
 GTGAAATTTAAAGACATTTGATTTCCGATGTAAAGGATTTTCTCATGAGTACAATAAGTCAATCAGTGTG
 CTCAAACTGCTTTATCTTATAAACAGCCATCTTAAATAGCAACGATTTGTGAGTACTGATATGATATATAATAA
 AATTAACAAGGAAAA

FIGURE 181

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52196
><subunit 1 of 1, 229 aa, 1 stop
><MW: 26017, pI: 4.73, NX(S/T): 0
MGDKIWLPPFVLLLAALPPVLLPGAAGFTPSLDSDFTLPAQKQKCFYQPMPLKASLEIEY
QVLGDAGLDIDFHLASPEGKTLVFEQRKSDGVHTVETEVGDYMFCDNTFSTISEKVIFEL
ILDNMGEQAQEDEDWKKYITGTDILDMKLEIDILESINSIKSRLSKSGHIQILLRAFEARDN
IQESNFDRVNFWSMVNLVVMVVVSAIQVYMLKSLFEDKRKSRT

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 195-217

N-myristoylation site.

amino acids 43-48

Tyrosine kinase phosphorylation site.

amino acids 55-62

FIGURE 182

CCATCCCTGAGATCTTTTATAAAAAACCCAGTCTTTGCTGACCAGACAAAGCATACCAGAT
CTCACCAGAGAGTCGCAGACACT**ATG**CTGCCTCCCATGGCCCTGCCAGTGTGTCCTGGATG
CTGCTTTCTGCCTCATTCTCCTGTGTGTCAGGTTCAAGGTGAAGAAAACCCAGAAGGAACTGCC
CTCTCCACGGATCAGCTGTCCCAAAGGCTCCAAGGCCTATGGCTCCCCCTGCTATGCCTTGT
TTTTGTCAACAAAATCCTGGATGGATGCAGATCTGGCTTGCCAGAAGCGGCCCTCTGGAAAA
CTGGTGTCTGTGCTCAGTGGGGCTGAGGGATCCTTCGTGTCTCCTGGTGAGGAGCATTAG
TAACAGCTACTCATACATCTGGATTGGGCTCCATGACCCACACAGGGCTCTGAGCCTGATG
GAGATGGATGGGAGTGGAGTAGCACTGATGTGATGAATTACTTTGCATGGGAGAAAAATCCC
TCCACCATCTTAAACCCCTGGCCACTGTGGGAGCCTGTCAAGAAGCACAGGATTTCTGAAGTG
GAAAGATTATAACTGTGATGCAAAGTTACCCATATGCTGCAAGTTCAAGGAC**TAG**GGCAGGT
GGGAAGTCAGCAGCCTCAGCTTGGCGTGCAGCTCATCATGGACATGAGACCAGTGTGAAGAC
TCACCTGGAAGAGAATATTCTCCCCAACTGCCCTACCTGACTACCTTGTCTATGATCCTCC
TTCTTTTTCTTTTTCTTCACCTTCATTTCAGGCTTTTCTCTGTCTTCCATGTCTTGAGATC
TCAGAGAATAATAAAAAATGTTACTTTATAAAAAAAAAAAAAAAAAAAAAA

FIGURE 183

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56965

<subunit 1 of 1, 175 aa, 1 stop

<MW: 19330, pI: 7.25, NX(S/T): 1

MLPPMALPSVSWMLLSCLILLCQVQGEETQKELPSPRISCPKGSKAYGSPCYALFLSPKSWM
DADLACQKRPSGKLVSVLSGAEGSFVSSLVRSISNSYSYIWIWGLHDPTQGSEPDGDGWEWSS
TDVMNYFAWEKNPSTILNPGHCGSLSRSTGFLKWKDYNCDAKLPYVCKFKD

Important features:

Signal peptide:

amino acids 1-26

C-type lectin domain signature.

amino acids 146-171

FIGURE 184

CCAGTCTGTGCGCCACCTCACTTGGTGTCTGCTGTCCCGCCAGGCAAGCCTGGGGTGAGAGC
ACAGAGGAGTGGGCCGGGACCATGCGGGGGACGCGGCTGGCGCTCCTGGCGCTGGTGTCTGGC
TGCCTGCGGAGAGCTGGCGCGGCCCTGCGCTGCTACGTCTGTCCGGAGCCACAGGAGTGT
CGGACTGTGTCACCATCGCCACCTGCACCACCAACGAAACCATGTGCAAGACCACACTCTAC
TCCCGGGAGATAGTGTACCCCTTCCAGGGGGACTCCACGGTGACCAAGTCCTGTGCCAGCAA
GTGTAAGCCCTCGGATGTGGATGGCATCGGCCAGACCCTGCCCCGTCTCTGCTGCAATACTG
AGCTGTGCAATGTAGACGGGGCGCCCGCTCTGAACAGCCTCCACTGCGGGGCCCTCACGCTC
CTCCCACTCTTGAGCCTCCGACTGTAGAGTCCCCGCCACCCCCATGGCCCTATGCGGCCCA
GCCCCGAATGCCTTGAAGAAGTGCCCCCTGCACCAGGAAAAAAAAAAAAAAAAAAAA

35
40
45
50
55
60
65
70
75
80
85
90
95
100

FIGURE 185

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56405

<subunit 1 of 1, 125 aa, 1 stop

<MW: 13115, pI: 5.90, NX(S/T): 1

MRGTRLALLALVLAACGELAPALRCYVCPEPTGVSDCVTIATCTTNETMCKTTLYSREIVYP
FQGDSTVTKSCASKCKPSDVGIGQTLVPVSCNTELCNVDGAPALNSLHCGALTLLPLLSLRL

Important features:

Signal peptide:

amino acids 1-17

N-glycosylation site.

amino acids 46-49

1000
900
800
700
600
500
400
300
200
100
0
100
200
300
400
500
600
700
800
900
1000

FIGURE 186

CTGCAGTCAGGACTCTGGGACCGCAGGGGGCTCCCGGACCCTGACTCTGCAGCCGAACCGGC
ACGGTTTCGTGGGGACCCAGGCTTGCAAAGTGACGGTCATTTTCTCTTTCTCCCTCTT
GAGTCCTTCTGAGATGATGGCTCTGGGCGCAGCGGGAGCTACCCGGGTCTTTGTGCGGATGG
TAGCGGCGGCTCTCGGCGGCCACCCTCTGCTGGGAGTGAGCGCCACCTTGAACTCGGTTCTC
AATTCCAACGCTATCAAGAACCTGCCCCACCGCTGGGCGGCGCTGCGGGGACCCAGGCTC
TGCAGTCAGCGCCGCGCGGGAATCCTGTACCCGGGCGGGAATAAGTACCAGACCATTGACA
ACTACCAGCCGTACCCGTGCGCAGAGGACGAGGAGTGCGGCACTGATGAGTACTGCGCTAGT
CCCACCCGCGAGGGGACGCGAGGCGTGCAAATCTGTCTCGCCTGCAGGAAGCGCCGAAACG
CTGCATGCGTCACGCTATGTGCTGCCCGGGAATTACTGCAAAAATGGAATATGTGTGTCTT
CTGATCAAAATCATTTCCGAGGAGAAATTGAGGAAACCATCACTGAAAGCTTTGGTAATGAT
CATAGCACCTTGATGGGTATTCCAGAAGAACCCCTTGTCTTCAAAAATGTATCACACCAA
AGGACAAGAAGGTTCTGTTTGTCTCCGGTCATCAGACTGTGCCTCAGGATTGTGTTGTGCTA
GACACTTCTGGTCCAAGATCTGTAAACCTGTCTGAAAGAAGGTCAAGTGTGTACCAAGCAT
AGGAGAAAAGGCTCTCATGGACTAGAAATATCCAGCGTTGTTACTGTGGAGAAGGTCTGTC
TTGCCGATACAGAAAGATCACCATCAAGCCAGTAATTCTTCTAGGCTTCACACTTGTGAGA
GACACTAAACCAGCTATCCAAATGCAGTGAACCTCTTTATATAATAGATGCTATGAAACC
TTTTATGACCTTCATCAACTCAATCCTAAGGATATACAAGTTCTGTGGTTTCAGTTAAGCAT
TCCAATAACACCTTCAAAAACCTGGAGTGTAAGAGCTTTGTTTCTTTATGGAACCTCCCCTG
TGATTGCAGTAAATTACTGTATTGTAAATTCTCAGTGTGGCACTTACCTGTAATGCAATGA
AACTTTTAATTATTTTCTAAAGGTGCTGCAGTCGCTATTTTCTCTTGTATGTAAATTT
TTGTACACATTGATTGTTATCTTGACTGACAAATATTCTATATTGAACTGAAGTAAATCAT
TCAGCTTATAGTTCTTAAAGCATAACCCCTTTACCCCATTTAATTCTAGAGTCTAGAACGCA
AGGATCTCTTGAATGACAAATGATAGGTACCTAAATGTAAACATGAAAACTAGCTTATT
TTCTGAAATGTACTATCTTAATGCTTAAATTATATTTCCCTTTAGGCTGTGATAGTTTTGA
ATAAAATTTAACATTTAAAAA

FIGURE 187

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57530

<subunit 1 of 1, 266 aa, 1 stop

<MW: 28672, pI: 8.85, NX(S/T): 1

MMALGAAGATRVFVAMVAAALGGHPLLGVSATLNSVLNSNAIKNLPPLGGAAGHPGSAVSA
APGILYPGGNKYQTTIDNYQPYPCAEDDECGTDEYCASPTRGGDAGVQICLACRKRRCMRH
AMCCPGNYCKNGICVSSDQNHFRGEIEETITESFGNDHSTLDGYSRRTTLSSKMYHTKGQEG
SVCLRSSDCASGLCCARHFWSKICKPVLKEGVCTKHRKKGSHGLEIFQRCYCGEGLSCRIQ
KDDHQAASNSSRLHTCQRH

Important features:

Signal peptide:

amino acids 1-23

N-glycosylation site.

amino acids 256-259

Fungal Zn(2)-Cys(6) binuclear cluster domain

amino acids 110-126

FIGURE 188

TGTGTTTCCCTGCAGTCAGAAATTTGGGACNGCAGGGGTTCCCGGACCTGATTTTGCAGCGGA
ACGGGAAGGTTTTGTGGGACCCAGGTTGAAATGACGGTCATTTTTTTCTTTCTCCTTCNG
GAGTCCTTNTGAGANGATGGTTTTGGGCGCAGCGGGAGCTAACCCGGTTTTTTGTNGCGATG
GTAGCGGCGGTTTTTCGGCGGCCACCTTNTGCTGGGAGTGAGCGCCACCTTGAATCGGTTTTTC
AATTCACGNTATCAAGAACCTGCCCCACCGNTGGGCGGCGCTGCGGGGCACCCAGGNTT
TGCAGTCAGCGCCGCGCCGGAATCCTGTACCCGGGCGGGAATAAGTACCAGACCATTGACA
ATTACCAGCCGTACCCGTGCGCAGAGGACGAGGAGTGCGGCACTGATGAGTACTGCGCTAGT
CCCACCCGCGGAGGGGANGCGGGCGTGCAAATNTGTNTNGCCTGCAGGAAGCGCCGAAAACG
CTGCATGCGTCANGCTATGTGCTGCCCCGGAATTACTGCAAAAATGGAATATGTGTGNTT
CTGATCAAAATCATTTCCGAGGAGAAATTGAGGAAACCATCACTGAAAGCTTTGGTAATGAT
CATAGCACCTTGGATGGG

FIGURE 189

GAGGAACCTACCGGTACCGGCCGCGCGCTGGTAGTCGCGGTGTGGCTGCACCTTACCAATCCCGTGCGCCGCGG
 CTGGGCGCTCGGAGAGTGCCTGTCTCTCTCTGCAACGCGGTGCTTGGGCTCGGCCAGGCGGGGCTCGCGGCCA
 GGGTTTGAGGATGGGGAGTAGCTACAGGAAGCGACCCCGCATGGCAAGATATATTTTGGGAATGAAAGGA
 AGTATTAGAAATAGCTGAGACCAATTCACAGATTAAATTTTGGGGACAGATTTGTGATGCTTGATTCACCTT
 TGAAGTAATGTAGACAGAAGTTCTCAAATTTGCATTTACATCAACTGGACAGAGTGAATCTTAATGTTTAC
 TTAATACAGAACTTGCATAAGAAAGAGATGGGAGTCTGGTTAAATAAAGATGACATATACAGAGACTGAAAG
 GATCAITCTCTGTTTCTGATAGTGTATGGCCATTTTAGTGGGCAGATCAGGATTTTACAGTTTACTTTGG
 AGTGCACAAACTCGCAAGCAGTAGAGAAATAAGACAAGCTTTCAGAAATTTGGCAATTTGAAGTTACATCTCGTAA
 AAACCCGAATAACCCAAATGCTACATGGCGATTTTAAATAATAAGACATTAAGTACATCAAGATGAGA
 TCTACGGAAAGATATGACAAATTTAGGAAAGGAGCTTCAGGATAATCAAGTGGCCAGTATGAAAGCTGAA
 CTATTATCGTATGATTTTGGTATTTATGATGATCTTGAATCATAACTTTGAAAGAGGAATTTGATGC
 TGCTGTAAATCTCGAGAACTGTGGTTTGTAAATTTTACTCCCCAGGCTGTTCACAGCTCGATGATTTAGTCC
 CACATGGAGAGACTTTGCTAAGAAAGTGGATGGTTACTTCGAATTTGAGCTGTAACTGTGGTGTATGATAGAAT
 GCTTTGCCGAATGAAAGGAGTCAACAGCTATCCAGTCTCTCAITTTTCGCTCGGAATGGCCCCAGTAAATA
 TCATGAGACAGATCAAGAGGAGTTTAGTGAATTTTGCATCGCAGCATGTAGAAAGTACAGTACAGAACTTTG
 GACAGGAAATTTTGTCAACTACACAACTGCTTTTGTGCTGGTATTTGGCTGGCTGATCACTTTTGTCTCAA
 AGGAGGAGATTTTGTGACTTCACAGACAGCATCAGGCTTAGTGGCATGTTGTTTCTCAACTCATTTGGATGTCAA
 AGAAATATATTTGGAAGTAAACATATAATCTTCAGATTTTGAACATCTTCGGCAACACACATAGAGGATCGTTT
 GGCTCATCATCGGTGGCTGTATTTTTCATTTTGGAAAAAATGAAATTCAAATAGTCTTGAGCTGAAAAAATC
 AAAAACTCTACTTAAAAATGATCATATTCAGTTGGCAGGTTTGACTGTTCCCTCGCACAGACATCTGTAGTAA
 TCTGTATGTTTTTCAGCGCTCTCAGCAGTATTTAAAGGACAAGGAACCAAGAAATGAAATTCATCATGAA
 AAGATTTCTATAGATATCTTGGCTTTGCAAGAAAGGTGAATCTCATGTTTACCAAGCTTGGACCTCAAAA
 TTTTCTCCCAATGACAAAGAACCATGGCTGTTGATTTCTTTGCCCTGGTGTCCACATGTCTGAGCTTTACT
 ACCAGAGTTGCAAGAGCATCAAACTCTTTTATGCTCAGCTTAAGTTTGGTACACTAGATTTGACAGTTTCA
 GGGACTCTGTAACTGTATAACATTCAGGCTTATCCACACACAGTGTATTCACACAGTCCAAACATCTCATAGTA
 TGAAGGACATCACTCTGCTGAACAAATCTTGGAGTTTATAGAGGATCTTATGAATCTTCAGTGGTCTCCCTTAC
 ACCACCACCTTCAACGAACCTAGTTACACAAAGAAACACAAAGAGTCTGGAATGTTGATTTCTATCTCCGT
 GTGTCTCTTCCGAAGTCTTAATGCCAGAAATGGAAGAAAGTGGCCGGACATTAATCTGAGCTGATCAACGTGGG
 CAGTATAGATTTGCCAACAGTATCATCTTTTGTGCCAGGAAACGTTCAAAGATACCTTGAGATAAGATTTT
 TCCCCAAATCAAATAAAGCTTATCAGTATCAGATTACAATGGTTGGAATAGGATGCTTATTTCCCTGAGAAT
 CTGGGGTCTAGGATTTTACCTCAAGTATCCACAGATCTAACACCTCAGACTTTCAGTGAAAAAGTTCTACAGG
 GAAAAATCAITGGGTGATTGATTTCTATGCTCTTGGTGTGGACCTTGCCAGAAATTTTGGCTCCAGAAATTTGAGCT
 CTTGGCTAGGATGATTAAGAAAGAAAGTGAAGCTGGAAGAAAGTAGCTCTGAGGCTTATGCTCAGACATGCCAGAA
 AGCTGGGATCAGGCGTCTTCAACTCTGTTAAGTTTATTTTACGAAAGAGCAAGAGAAATTTTCAAGAAAGCA
 GATAAATACAGAGATGCAAAAGCAATCGCTGCCITTAATAGTGAATAATTTGAAATCTCCGAAATCAAGGCA
 GAGGAATAAGAGTGAATTTTGAATGTGAAGATGAAGAAAGTTTAAAGAAATTTGACAGATGACATCAG
 AAGACACTATTTAGAAATGTACATTTATGATGGGAATGAATGAACATTTATCTAGACTTCGAGTTGTAATGCCA
 GAATTTCAAGCAGCTGGTGTAAAGAGGGCTGCAAACTTTTCTGTAAGGGCGGGTTTATAAATATTTTA
 GACTTTGCGAGCTATAATATATGGTTCACACATGAGAACAGAAATAGAGTACATCATGATTTCTGTTATTTGGT
 TTTAACCAACTTTAAAAAAATATTAAGCATTTCTAGCTCAGGCCATACAAAGATGAGGCTGGATTCAGTCCAGT
 GACCATAGATTTGCTGTCCCTCGACGCACTTATAATGTTTTCAGGTGGCTGGCTTGAACATGAGCTGCTGTGGCT
 ATCTACATCAAGATGTCTAAGTTGTATAAGTCCACTTTCCCTTCAGGTTTTTGGCTGAGCTGAAAGAGGTAATCT
 TAGTTTGTGTCATGTTCTCTCAATAAAGCTATCCCTAACCATATATATTTATTTGTTTGTGTTTAAACACACCAT
 GATGTGGCAGAGTAAACCAACCTGTTATGCTGATATTTATGAGGAGATTTCTTATTTGTTTCTTCTCTCA
 AAGGTTGAAAAAATGCTTTAAATTTTTCAGGCCGAGAACAGTTCAGCAGATTTGAGTGTGACACAGTAAGTAC
 AATTTGAGCAACAGGTAAGTGCAAAATCTGTAGTTTGTCTGTATCATCCAGGAAAACTGAGGGAAGAAATTA
 TAGCAATTAACCTGAGATGTAGAGTATCTTAAATATGTTATCAAGTATTTAGAGTCTTATTTTAAAGATA
 TGTGTTTCATGTTTCTGAAATGGCTTTCATAGAAATTTTCCCATGAGTAGTGTGATTTTGGAGCATCTAATAT
 TACATATTTTGGCTCTGAACTTTGTTTACCTGTATCTTATTTACATTTGGTTTCTTTTCTTATTTTGG
 TTTTCTCTCTCTGTCAGCTATTTATTTATTCAAATAGGAAAAATTTACTTTACAGGTTGTTTACTGTAGCTTAT
 AATGATCTGTAGTTATTTCCAGTTTACTAGTTTACTGTCAGAGGCGTCCCTTTTTCAGATAAATTTGACATAA
 ACTGAAGTTATTTTATAAGAAATCAAGTATATAAATCTAGGAAAGGAGTCTTCTAGTTTCTGTTGTTTGA
 CTTCAAGAAATCAAAATTTGTCAGTAACTGTAGTTTGTAGTTATATATTTACAGTGTACAGATGTTAAAAAT
 CCAATCAGTCAAAGAGGTCAATGAATTAAGAGCTTGCAACTTTTCAAAAAAATTTTAAAAAT

FIGURE 190

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56439
<subunit 1 of 1, 747 aa, 1 stop
<MW: 86127, pI: 7.46, NX(S/T): 2
MGVWLNKDDYIRDLKRIILCFLIVYMAILVGTQDQDFYSLLGVSKTASSREIRQAFKKLALKL
HPDKNPNNPNAHGDFLKNRAYEVLKDEDLRKKYDKYGEKLEDNQGGQYESWNYYRYDFGI
YDDDEPIITLERREFDAAVNSGELWFVNFYSPGCSHCHDLAPTWRDFAKEVDGLLRIGAVNC
GDDRMLCRMKGVNSYPSLFI FRSGMAPVKYHGD RSKE SLVSFAMQHVRSTVTELWTGNFVNS
IQTAF AAGIGWLITFC SKGGDCLTSQTRLRLSGMLFLNSLDAKEIYLEVIHNLDPDFELLSAN
TLEDRLAHRWLLFFHFGKNENSNDPELKKLKTLLKNDHIQVGRFDCSSAPDICSNLYVFQP
SLAVFKGQGTKEYEIIHGKKILYDILAFAKESVNSHVTTLG PQNFPANDKEPWLVDFFAPWC
PPCRALLPELRRASNLLYGQLKFGTLDCTVHEGLCNMYNIQAYPTTVVFNQSNIEHEYEGHHS
AEQILEFIEDLMNPVS VSLTPTTFNELVTQRKHNEVWMVDFYSPWCHPCQVLMPEWKRMART
LTGLINVGSIDCQQYHSFCAQENVQRYPEIRFFPPKSNKAYQYHSYNGWNRDAYSLRIWGLG
FLPQVSTDLTPQTFSEKVLQGNHWVIDFYAPWCGPCQNFAPFELLARMIKGKVKAGKVDC
QAYAQTCQKAGIRAYPTVKFYFYERAKRNFQEEQINTRDAKAI AALISEKLETLRNQGKRNKDEL

Important features:

Endoplasmic reticulum targeting sequence.

amino acids 744-747

Cytochrome c family heme-binding site signature.

amino acids 158-163

Nt-dnaJ domain signature.

amino acids 77-96

N-glycosylation site.

amino acids 484-487

FIGURE 191

AGACAGTACCTCCTCCCTAGGACTACACAAGGACTGAACCAGAAGGAAGGACAGAGCAAA
GCCATGAAACATCATCCTAGAAATCCTTCTGCTTCTGATCACCATCATCTACTCCTACTTGGA
GTCGTTGGTGAAGTTTTTTCATTCTCAGAGGAGAAAACTGTGGCTGGGGAGATTGTTCTCA
TTACTGGAGCTGGGCATGGAATAGGCAGGCAGACTACTTATGAATTTGCAAAACGACAGAGC
ATATTGGTTCTGTGGGATATTAATAAGCGCGGTGTGGAGGAACTGCAGCTGAGTGCCGAAA
ACTAGGCGCTCACTGCGCATGCGTATGTGGTAGACTGCAGCAACAGAGAAGAGATCTATCGCT
CTCTAAATCAGGTGAAGAAAGAAGTGGGTGATGTAACAATCGTGGTGAATAATGCTGGGACA
GTATATCCAGCCGATCTTCTCAGCACCAAGGATGAAGAGATTACCAAGACATTTGAGGTCAA
CATCTAGGACATTTTTGGATCACAAAAGCACTTCTTCCATCGATGATGGAGAGAAATCATG
GCCACATCGTCACAGTGGCTTCAGTGTGCGGCCACGAAGGGATTCCCTTACCTCATCCCATAT
TGTTCCAGCAAATTTGCGCGTGTGTGGCTTTCACAGAGGTCTGACATCAGAACCTTCAGGCCCTT
GGGAAAACTGGTATCAAAACCTCATGTCTCTGCCCAGTTTTTGTGAATACTGGGTTCACCA
AAAATCCAAGCACAAGATTATGGCCTGTATTGGAGACAGATGAAGTCGTAAGAAGTCTGATA
GATGGAATACTTACCAATAAGAAAAATGATTTTTGTTCCATCGTATATCAATATCTTTCTGAG
ACTACAGAAGTTTCTTCTGAACGCGCCTCAGCGATTTTAAATCGTATGCAGAAATTTCAAT
TTGAAGCAGTGGTTGGCCACAAAATCAAAATGAAATTGAATAAATAAGCTCCAGCCAGAGATG
TATGCATGATAATGATATGAATAGTTTCGAATCAATGCTGCAAAGCTTTATTTACATTTTTT
TCAGTCTCTGATAATATTAAAAACATTGGTTTGGCACTAGCAGCAGTCAACGAACAAGATTA
ATTACCTGTCTTCTCTGTTTCTCAAGAATATTTACGTAGTTTTTCATAGGTCTGTTTTTCTCTT
TCATGCCTCTTAAAAACTTCTGTGCTTACATAAACATACTTAAAAAGGTTTTCTTTAAGATAT
TTTTTTTTTCCATTTAAAGGTGGACAAAAGCTACCTCCCTAAAAGTAAATACAAGAGAAT
TATTTACACAGGGAAGGTTTAAAGACTGTTCAAGTAGCATTCCAATCTGTAGCCATGCCACAG
AATATCAACAAGAACACAGAATGAGTGCACAGCTAAGAGATCAAGTTTCAGCAGGCAGCTTT
ATCTCAACCTGGACATATTTTAAAGATTGAGCATTGAAAGATTTCCCTAGCCTCTTCTCTTT
TCATTAGCCCAAAACGGTGCAACTCTATTCTGGACTTTATTACTTGATTCTGTCTTCTGTAT
AACTCTGAAGTCCACAAAAGTGGACCTCTATATTTCCTCCCTTTTTATAGTCTTTATAAGA
TACATTATGAAAGGTGACCGACTCTATTTTAAATCTCAGAATTTTAAAGTTCTAGCCCCATGA
TAACCTTTTTCTTTGTAATTTATGCTTTTCATATATCCTTGGTCCCAGAGATGTTTAGACAAAT
TTTAGGCTCAAAAATTAAGACTAACACAGGAAAAGGAACGTACTGGCTATTACATAAGAA
CAATGGACCCAAGAGAAGAA

FIGURE 192

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56409

<subunit 1 of 1, 300 aa, 1 stop

<MW: 33655, pI: 9.31, NX(S/T): 1

MNIILEILLLLITIIYSYLESLVKFFIPQRRKSVAGEIVLITGAGHGIGRQTTFEFAKRQSI
LVLWDINKRGVEETAACRKLGVTAHAYVVDCSNREEIYRSLNQVKKEVGDTVIVVNNAGTV
YPADLLSTKDEEITKTFEVNI LGHFWITKALLPSMMERNHGHIVTVASVCGHEGIPYLIPYC
SSKFAAVGFGHRLTSELQALGKTGIKTSCLCPVFVNTGFTKNPSTRLWPVLETDEVVRSLID
GILTNNKMIFVPSYINIFLRLQKFLPERASAILNRMQNIQFEAVVGHKIKMK

Important features:

Signal peptide:

amino acids 1-19

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-33 and 58-61

Short-chain alcohol dehydrogenase family protein

amino acids 165-202, 37-49, 112-122 and 210-219

FIGURE 193

CGGCGGCGGCTGCGGGCGCGAGGTGAGGGGCGCGAGGTGAGGGGCGCGAGGTGCCAGCAGG
ATGCCCCGGCTCTGCAGGAAGCTGAAGTGAGAGGCCCGAGAGGGCCAGCCCGCCCGGGGCG
AGGATGACCAAGGCCCGGCTGTTCCGGCTGTGGCTGGTGCTGGGGTCGGTGTTTCATGATCCT
GCTGATCATCGTGTA¹CTGGGACAGCGCAGGCGCGCGCACTTCTACTTGCACACGTCCTTCT
CTAGGCCGCACACGGGGCCGCCGCTGCCCACGCCCGGGCCGACAGGGACAGGGAGCTCACG
GCCGACTCCGATGTCTGACGAGTTTCTGGACAAGTTTCTCAGTGCTGGCGTGAAAGCAGAGCGA
CCTTCCCAGAAAGGAGACGAGCAGCCGCTGCGCCGGGGAGCATGGAGGAGAGCGTGAGAG
GCTACGACTGGTCCCCGCGCAGCCCGGCGCAGCCAGACCAGGGCCGGCAGCAGGCGGAG
CGGAGGAGCGTGCTGCGGGGCTTCTGCGCCAAC²TCCAGCCTGGCCTTCCCCACCAAGGAGCG
CGCATTCGACGACATCCCCAACTCGGAGCTGAGCCACCTGATCGTGGACAGCCGGCACGGGG
CCATCTACTGCTACGTGCCCAAGGTGGCCTGCACCAACTGGAAGCGCGTGATGATCGTGCTG
AGCGGAAGCCTGCTGCACCGCGGTGCGCCCTACCGCGACCCGCTGCGCATCCCGCGCGAGCA
CGTGCAACAACGCCAGCGCGCACCTGACCTTCAACAAGTTCTGGCGCCGCTACGGGAAGCTCT
CCCGCCACCTCATGAAGTCAAGCTCAAGAAGTACACCAAGTTCTCTTCGTGCGCGACCCC
TTCGTGCGCCTGATCTCCGCCTTCCGCAGCAAGTTCGAGCTGGAGAACGAGGAGTTCTACCG
CAAGTTCGCGGTGCCCCATGCTGCGGCTGTACGCCAACCAACCAGCCTGCCGCCCTCGGCGC
GCGAGGCCTTCCGCGCTGGCCTCAAGGTGTCTTCGCCAACTTCATCCAGTACCTGCTGGAC
CCGCACACGGAGAAGCTGGCGCCCTTCAACGAGCACTGGCGGCAGGTGTACCGCTCTGCCA
CCCGTGCCAGATCGACTACGACTTCTGTTGGGAAGCTGGAGACTCTGGACGAGGACGCGCGC
AGCTGCTGCAGCTACTCCAGGTGGACCGGCAGCTCCGCTTCCCCCGAGCTACCGGAACAGG
ACCGCCAGCAGCTGGGAGGAGGACTGGTTGCGCAAGATCCCCCTGGCCTGGAGGCAGCAGCT
GTATAAACTCTACGAGGCCGACTTTGTTCTCTTCGGCTACCCCAAGCCCGAAAACTCCTCC
GAGACTGAAAGCTTTGCGTGTGCTTTTCTCGCGTGCTGGAACCTGACGCACGCGCACTCC
AGTTTTTTTATGACCTACGATTTTGCAATCTGGGCTTCTTGTTCACTCCACTGCCTCTATCC
ATTGAGTACTGTATCGATA³TGTTTTTTAAGATTAATATATTTTCAGGTATTTAATACGA

FIGURE 194

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56112

<subunit 1 of 1, 414 aa, 1 stop

<MW: 48414, pI: 9.54, NX(S/T): 4

MTKARLFRLWLVLGVSVMILLIIVYWDSAGAAHFYLHTSFSPHTGPPLTPGPDRDRELTA
DSDVDEFLDKFLSAGVKQSDLPRKETEQPPAPGSMEE SVRGYDWSPRDARRSPDQGRQQAER
RSVLRGFCANSSLAFFTKEAFDDIPNSELSHLIVDDRHGAIYCYVPKVACTNWKVMIVLS
GSLLRHGAPYRDLPLRIPREHVHNASAHLT FNKFWRRYGKLSRHLMKVKLKKYTKFLFVRDPF
VRLISAFRSKFELNEEFYRKFAVPMRLRYANHTSLPASAREAFRAGLKVSFANFIQYLLDP
HTEKLAPFNEHWRQVYRLCHPCQIDYDFVGKLETLDEDAQLQLQVDRQLRFPPSYRNRT
ASSWEEDWFAKIPLAWRQQLYKLYEADFVLFGYPKPENLLRD

Important features:

Signal peptide:

amino acids 1-31

N-glycosylation sites.

amino acids 134-137, 209-212, 280-283 and 370-373

TNFR/NGFR family cysteine-rich region protein

amino acids 329-332

FIGURE 195

TCGGGCCAGAATTTCGGCACGAGGCGGCACGAGGGCGACGGCCTCACGGGGCTTTGGAGGTGA
AAGAGGCCCAGAGTAGAGAGAGAGAGACCGACGTACACGGGATGGCTACGGGAACGCGCT
ATGCCGGGAAGGTGGTGGTCTGTGACCGGGGGCGGGCGCGCATCGGAGCTGGGATCGTGCGC
GCCTTCGTGAACAGCGGGGCCCGAGTGGTTATCTGCGACAAGGATGAGTCTGGGGGCCGGGC
CCTGGAGCAGGAGCTCCCTGGAGCTGTCTTTATCCTCTGTGATGTGACTCAGGAAGATGATG
TGAAGACCCTGGTTTCTGAGACCATCCGCCGATTTGGCCGCCTGGATTGTGTTGTCAACAAC
GCTGGCCACCACCCACCCCCACAGAGGCCTGAGGAGACCTCTGCCCAGGGATTCCGCCAGCT
GCTGGAGCTGAACCTACTGGGGACGTACACCTTGACCAAGCTCGCCCTCCCTACCTGCGGA
AGAGTCAAGGGAATGTCATCAACATCTCCAGCCTGGTGGGGGCAATCGGCCAGGCCCAGGCA
GTTCCCTATGTGGCCACCAAGGGGGCAGTAACAGCCATGACCAAGCTTTGGCCCTGGATGA
AAGTCCATATGGTGTCCGAGTCAACTGTATCTCCCAGGAAACATCTGGACCCCGTGTGGG
AGGAGCTGGCAGCCTTAATGCCAGACCCCTAGGGCCACAATCCGAGAGGGCATGCTGGCCCAG
CCACTGGGCCGCATGGGCCAGCCCGCTGAGGTGCGGGCTGCGGCAGTGTTCCTGGCCTCCGA
AGCCAACTTCTGCACGGGCATTGAACTGCTCGTGACGGGGGGTGACAGCTGGGGTACGGGT
GCAAGGCCAGTCGGAGCACCCCGTGGACGCCCCGATATCCCTTCCTGATTTCTCTCATTT
CTACTTGGGGCCCCCTTCTTAGGACTCTCCACCCCAAACCTCCAACCTGTATCAGATGCAGC
CCCCAAGCCCTTAGACTCTAAGCCCAGTTAGCAAGGTGCCGGGTACCCCTGCAGGTTCCCAT
AAAAACGATTTGCAGCC

FIGURE 196

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56045

<subunit 1 of 1, 270 aa, 1 stop

<MW: 28317, pI: 6.00, NX(S/T): 1

MATGTRYAGKVVVVTGGGRGIGAGIVRAFN SGARVVICDKDES GGRALEQELPGAVFILCD
VTQEDDVKTLVSETIRRFGRLCDV VNNAGHHPPPQRPEETSAQGFRQLLELNLLGTYTLTKL
ALPYLRKSQGNVINISSLVGAIGQAQAVPYVATKGAVTAMTKALALDESPYGVRVNCISPGN
IWTPLWEEL AALMPDPRATIREGMLAQPLGRMGQPAEVGAAAVFLASEANFCTGIELLVTTGG
AELGYGCKASRSTPVDAPDIPS

Important features:

N-glycosylation site.

amino acids 138-141

Short-chain alcohol dehydrogenase family protein

amino acids 10-22, 81-91, 134-171 and 176-185

FIGURE 197

AGGCGGGCAGCAGCTGCAGGCTGACCTTGACAGCTTGGCGGAAATGACTGGCCTCACAACCTG
CTGTTTTCTTCTTACCATTTCATCTTCTGGGGCTGGGGCCAGCCCAGGAGCCCCAAAAGCAA
GAGGAAGGGGCAAGGGCGGCTGGGCCCCCTGGCCCCCTGGCCCTCACCAGGTGCCACTGGACC
TGGTGTCACGGATGAAACCGTATGCCCGCATGGAGGAGTATGAGAGGAACATCGAGGAGATG
GTGGCCCGCTGAGGAACAGCTCAGAGCTGGCCAGAGAAAGTGTGAGGTCAACTTGACAGCT
GTGGATGTCCAACAAGAGGAGCCTGTCTCCCTGGGGCTACAGCATCAACCACGACCCAGCC
GTATCCCCGTGGACCTGCCGGAGGCACGGTGCTGTGTCTGGGCTGTGTGAACCCCTTCACC
ATGCAGGAGGACCGCAGCATGGTGAGCGTGCCGGTGTTCAGCCAGGTTCCCTGTGCGCCGCCG
CCTCTGCCCCCACC GCCCCGACAGGGCCTTGCCGCCAGCGCGCAGTCATGGAGACCATCG
CTGTGGGCTGCACCTGCATCTTCTTGAATCACCTGGCCAGAAGCCAGGCCAGAGCCCGAGA
CCATCCTCCTTGACACCTTTGTGCCAAGAAAGGCCTATGAAAAGTAAACACTGACTTTTGAAA
GCAAG

FIGURE 198

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59294

<subunit 1 of 1, 180 aa, 1 stop

<MW: 20437, pI: 9.58, NX(S/T): 1

MDWPHNLLFLLTISIFLGLGQPRSPKSKRKGGQGRPGPLAPGPHQVPLDLVSRMKPYARMEEY
ERNIEEMVAQLRNSSELAQRKCEVNLQLWMSNKRSLSPWGYSINHDPRI PVDLPEARCLCL
GCVNPFTMQEDRSMVSVPVFSQVPVRRRLCPPPPRTGPCRQRAVMETI AVGCTCIF

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 75-78

Homologous region to IL-17

amino acids 96-180.

FIGURE 199

GCGCCGCCAGGCGTAGGCGGGGTGGCCCTTGCCTCTCCCGCTTCTTGA AAAACCCGGCGGG
CGAGCCGAGGCTGCGGGCCGGCCGCTGCCCTTCCCCACACTCCCCGCCGAGAAGCCTCGCTCG
GCGCCCAACATGCGGGGTGGGCGCTGCGGCCCGCAGCTAACGGCGCTCCTGGCCGCTTGGAT
CGCGGCTGTGGCGGCGACGGCGAGGCCCGAGGAGGCGCGCTGCGCCCGGAGCAGAGCCGGG
TCCAGCCCATGACCGCCTCCAACCTGGACGCTGGTGATGGAGGGCGCAGTGATGCTGAAATTT
TACGCCCCATGGTGTCCATCCTGCCAGCAGACTGATTCAGAATGGGAGGCTTTTGC AAAGAA
TGGTGAAATACTTCAGATCAGTGTGGGGAAGGTAGATGTCATTCAAGAACCAAGTTTGAGTG
GCCGCTTCTTTGTCAACCACTCCTCCAGCATT TTTTCATGCAAAAGGATGGGATATTTCCGCCGT
TATCCTGGCCCAAGGAATCTTGAAGACCTTGCAGAATATATCTTTAGAGAGAAGAAATGGCAATC
AGTCGAGCCTCTGACTGGCTGGAAATCCCCAGCTTCTCTAACGATGCTCTGGAATGGCTGGTC
TTTTTAGCATCTCTGGCAAGATATGGCATCTTCACAACTATTTCCAGTGACTCTTGGAAAT
CCTGCTTGGTGTTCTTATGTGTTTTTTCGTCATAGCCACCTTGGTTTTTGGCCCTTTTTATGGG
TCTGCTCTGGTGGTAATATCAGAATGTTTCTATGTGCCACTTCCAAGGCATTATCTGAGC
GTTCTGAGCAGAAATCGGAGATCAGAGGAGGCTCATAGAGCTGAACAGTTGCAAGATGCGGAG
GAGGAAAAAGATGATTCAAATGAAGAAGAAAAACAAAGACAGCCTTGTAGATGATGAAGAA
GAAAGAAGATCTTGGCGATGAGGATGAAGCAGAGGAAGAAAGAGGAGGAGGACAACCTTGGCTG
CTGGTGTGGATGAGGAGAGAAGTGAGGCCAATGATCAGGGGCCCCCGAGGAGAGGACGGTGTG
ACCCGGGAGGAAGTAGAGCCTGAGGAGGCTGAAGAAGGCATCTCTGAGCAACCCTGCCAGC
TGACACAGAGGTGGTGAAGACTCCTTGAGGCAGCGTAAAGTCAGCATGCTGACAAGGGAC
TG TAGATTAAATGATGCGTTTTCAAGAATACACACCAAAACAATATGTGAGCTTCCCTTTGG
CCTGCAGTTTGTACCAAAATCCTTAATTTTTCTCTGAATGAGCAAGCTTCTCTTAAAGATGCT
CTCTAGTCATTTGGTCTCATGGCAGTAAGCCTCATGTATACTAAGGAGAGTCTTCCAGGTGT
GACAATCAGGATATAGAAAAACAAACGTAGTGTGGGATCTGTTTGGAGACTGGGATGGGAA
CAAGTTCACTTTACTTAGGGGTGAGAGAGTCTCGACCAGAGGAGGCCATTCCCAGTCTCTAATC
AGCACCTTCCAGAGACAAGGCTGCAGGCCCTGTGAAATGAAAGCCAAGCAGGAGCCTTGGCT
CCTGAGCATCCCCAAAGTGTAACGTAGAAGCCTTGCACTCCTTTTTCTTGTTGTAAGATTTTTAT
TTTTGTCAAATGCAAGAAACATCAGGCACCACAGTGCATGAAAAATCTTTACAGCTAGAA
ATTGAAAAGGCCCTTGGGTATAGAGAGCAGCTCAGAAGTCATCCAGCCCTCTGAATCTCCTG
TGCTATGTTTTATTTCTTACCTTTAATTTTTCCAGCATTTCCACCATGGGCATTCAAGCTCT
CCACACTTCTCACTATTATCTCTTGGTCAGAGGACTCCAATAACAGCCAGGTTTACATGAAC
TGTGTTTTGTTCACTCTGACCTAAGGGGTTTAGATAATCAGTAACCATAACCCCTGAAGCTGT
GACTGCCAAACATCTCAAATGAAATGTTGTGGCCATCAGAGACTCAAAGGAAGTAAGGATT
TTACAAGACAGATTA AAAAAAAAAATGTTTTGTGCCAAATATAGTTGTTGTGATTTTTTTTT
AAGTTTTCTAAGCAATATTTTTCAAGCCAGAAGTCTCTAAGTCTTGCCAGTACAAGGTAGT
CTTGTGAAGAAAAGTTGAATACTGTTTTGTTTTCTATCTCAAGGGGTTCCCTGGGTCTTGAAC
TACTTTAATAATAACTAAAAAACCACTCTGATTTTTCTTCAGTGATGTGCTTTTGGTGAAA
GAATTAATGAACCTCAGTACTGAAAGTGAAGATTTGATTTTTGTTTCCATCTCTGTAAATC
TTCCAAGAAATTATATCTTTGTAAATCTCTCAATACTCAATCTACTGTAAGTACCCAGGGAG
GCTAATTTCTTT

FIGURE 200

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56433

<subunit 1 of 1, 349 aa, 1 stop

<MW: 38952, pI: 4.34, NX(S/T): 1

MAGGRCGPQLTALLAAWIAAVAATAGPEEAAALPPEQSRVQPMASNWTLVMEGEWMLKFYAP
WCPSCQQTDSWEAFKNGEILQISVGKVDVIQEPGLSGRFFVTTLPAFFHAKDGIFFRRYRG
PGIFEDLQNYILEKKWQSVEPLTGWKSPASLTMSGMAGLFSISGKIWHLHNYFTVTLGIPAW
CSYVFFVIATLVFGLFMGLVLVVISFCFYVPLPRHLSERSEQNRRSEEAHRAEQLQDAEEEEK
DSDNEEENKDSLVDDEEEKEDLGDEDEAEDEEEEDNLAAGVDEERSEANDQGPPEGEDGVTRE
EVEPEEAEEGISEQPCPADTEVVEDSLRQRKSHADKGL

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 191-211

N-glycosylation site.

amino acids 46-49

Thioredoxin family proteins. (homologous region to disulfide isomerase)

amino acids 56-72

Flavodoxin proteins

amino acids 173-187

FIGURE 201

ATCTGGTTGAACTACTTAAAGCTTAATTTGTAAACTCCGGTAAGTACCTAGCCCACATGATT
TGACTCAGAGATTCTCTTTTGTCACAGACAGTCATCTCAGGGGCAGAAAGAAAGAGCTCC
CAAATGCTATATCTATTACAGGGCTCTCAAGAACAAATGGAATATCATCTGATTTAGAAAAAT
TTGGATGAAGATGGATATACCTCAATTACACTTCGACTCTCAAAGCAATACCAGGATAGCTGT
TGTTTTAGAGAAAGGATCGTGTGCTGCATCTCTCTCTGGCGCCCTCATGTGCTGAATTTTGG
GAATCCTATGCTTGGTAATACTGGTGATAGCTGTGGTCTGGGTACCATTGGGGGTTCTTTCC
AGCCCTTGCTCTCTCAATTGGATTATATATGAGAAGAGCTGTTATCTATTACAGCATGTCACT
AAATTCCTGGGATGGAAGTAAAGACAATGCTGGCAACTGGGCTCTAATCTCTCTAAAGATAG
ACAGCTCAAATGAATTTGGGATTTATAGTAAAAACAAGTGTCTTCCCAACCTGATAATCATTT
TGGATAGGCCCTTCTCGGCCCCAGACTGAGGTACCATTGGCTCTGGGAGGATGGATCAACATT
CTCTTCTAACTTATTTTACAGATCAGAACCCAGCTACCCAAAGAAAAACCATCTCCAAATTTGTG
TATGGATTACAGTGTGTCAGTCATTTATGACCAACTGTGTAGTGTGCCCTCATATAGTATTTGT
GAGAAGAAGTTTTCAATGTAAAGAGGAAGGTTGAGAAGGAGAGAGAAATATGTGAGGTAGTA
AGGAGGACAGAAAAACAGAACAGAAAAAGAGTAACAGCTGAGGTCAAGATAAATGCAGAAAAATG
TTTAGAGAGCTTGGCCCAACTGTAATCTTAAACAGAAATGAAGGGAGAGGCTGTGATTTCT
GTATTTGTGACACTACAGGTAGGCTAGTATTATTTTTCTAGTTAGTAGATCCCTAGACATGG
AATCAGGGCAGCCAAAGCTTGAGTTTTATTTTTTATTATTTATTTTTTTGAGATAGGGTCT
CACTTTGTTTACCAGGCTGGAGTGCAGTGGCACAATCTCGACTCACTGCAGCTATCTCTCGC
CTCAGCCCTCAAGTAGCTGGGACTACAGGTGCATGCCACCATGCCAGGCTAATTTTTGGTG
TTTTTTGTAGAGACTGGGTTTTGCCATGTTGACCAAGCTGGTCTCTAACTCCTGGGCTTAAG
TGATCTGCCCGCTTGGCCTCCCAAAGTGTGGGATTACAGATGTGAGCCACCACCTGGC
CCCAAGCTTGAATTTTCATTCTGCCATTGACTTGGCATTACCTTGGGTAAGCCATAAGCGA
ATCTTAATTTCTGGCTCTATCAGAGTTGTTTCATGCTCAACAATGCCATTGAAGTGACCGGT
GTGTTGCCACGATTTGACCCCAACTTCTAGCAGTATATCAGTTATGAACGAGGGTGAAT
ATATTTCTGAATAGCTAAATGAAGAAATGGGAAAAAATCTTCAACCAAGTCAGAGCAATTTT
ATTATTTTCATCAGTATGATCATAATTATGATTATCATCTTAGTAAAAAGCAGGAACCTCTA
CTTTTTCTTTATCAATTAATAGCTCAGAGAGTACATCTGCCATATCTCTAATAGAACTCTTT
TTTTTTTTTTTTTTTTTTTGGACAGAGTTCGCTCTTGTTGCCAGGCTGGAGTGCAACGG
CAGGATCTCGGCTCACCGCAACCTCCGCCCTTGGGTTCAAGCAATTTCTCTGCTCAGCCT
CCCAAGTAGCTGGGATTCAGTACAGGCACCACCAACCCGGCTAATTTTGTATTTTTTAGT
AGAGACAGGGTTTCTCCATGTGCGTCAAGGTAGTCCGCACTCCGACCTCAAGTGATCTGC
CTGCCCTCGGCTCCCAAGTGTGCGGATTACAGGCGTGAGCCACTGCACCCAGCCTAGAATCT
TGTATAATATGAATTTAGGGAACCTGCTCTCATAGGAAAGTTTCTGCTTTTAAATACA
AAAAACATAAAAAATACATAAAATCTGATGATGAATATAAAAAAGTAACCAACCTCATTGGA
ACAAGTATTAACATTTTGGAAATATGTTTTATTAGTTTTGTGATGTACTGTTTACAATTTTT
ACCATTTTTTTCAGTAATTACTGTAAAAATGGTATTATTGGAATGAACTATATTTCTCATG
TGCTGATTGTCTTATTTTTTTCATACTTTCCCACTGGTGCTATTTTATTTCATGAGATA
TTTCTGTTACTAGGGAGGCATTTACAGTCTCTAATGTTGATTAAATATGTGAAAAGAAAT
TGTACCAATTTTACTAAATTTATGCAGTTTAAATGGATGATTTTATGTTATGTGGATTTTCA
TTCAATAAAAAAACTCTTATCAAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 202

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53912

<subunit 1 of 1, 201 aa, 1 stop

<MW: 22563, pI: 4.87, NX(S/T): 1

MEYHPDLENLDEDGYTQLHFDSQSNTRIAVVSEKSCAASPPWRLLIAVILGILCLVILVIAV
VLGTMGVLSSPCPPNWIIYEKSCYLFMSLSNSWDGSKRQCWQLGSNLLKIDSSNELGFIVKQ
VSSQPDNSFWIGLSRPQTEVPWLWEDGSTFSSNLFQIRTTATQENPSPNCVWIHVSVIYDQL
CSVPSYSICEKKFSM

Important features:

Type II transmembrane domain:

amino acids 45-65

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 197-200

N-myristoylation sites.

amino acids 35-40 and 151-156

Homologous region to LDL receptor

amino acids 34-67 and 70-200.

[illegible][illegible]

FIGURE 204

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50921
<subunit 1 of 1, 693 aa, 1 stop
<MW: 77738, pI: 8.87, NX(S/T): 7
MTPQSLQLQTTLFLLSLLFLVQGAHGRGHREDFRFC SQRNQTHRSSLHYKPTPDLRISIENSE
EALT VHAPPPAAHPASRSFPDPGRGLYHFCLYWNRHAGRLHLLYGRDFFLLSDKASSLLCFQH
QEESLAQGPPLLATSVTSWWSPQNI SLPSAASFTFSFHSPPHTAAHNASVDMCELKRDQLL
SQFLKHPQKASRRRPSAAPASQQLQSLESKLT SVRFMGDMVSFEEDRINATVWKLQPTAGLQD
LHIHSRQEEEQSEIMEYSVLLPRTL FQRTKGRSGEAEKRLLLVDFSSQALFQDKNSSQVLGE
KVLGIVVQNTKVANLTEPVVLT FQHQLQPKNVTLCQVFWVEDPTLSSPGHWSSAGCETVRRE
TQTSCFCNHLTYFAVLMVSSVEVD AVHKHYLSLLSYVGCVVSAACLVITIAAYLC SRVPLPC
RRKPRDYTIKVHMNLLLA VFLLDTSFLLSEPVALTGSEAGCRASAI FLHFSLLTCLSWMGLE
GYNLYRLVVEVFGT YVPGYLLKLSAMGWGFPI FLVTLVALVDVDNYGPI ILAVHRTPEGVIY
PSMCWIRDSLVS YITNLGLPSLVFLFN MAMLATMVVQILRLRPHTQKWSHVLTLLGLSLVLG
LPWALIFFSFASGT FQLVVLYLFSIITS FQGFLIFIWYWSMRLQARGGPSPLKNSDSARLP
ISSGSTSSSRI
```

Important features:

Signal peptide:

amino acids 1-25

Putative transmembrane domains:

amino acids 382-398, 402-420, 445-468, 473-491, 519-537, 568-590
and 634-657

Microbodies C-terminal targeting signal.

amino acids 691-693

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 198-201 and 370-373

N-glycosylation sites.

amino acids 39-42, 148-151, 171-174, 234-237, 303-306, 324-327
and 341-344

G-protein coupled receptors family 2 proteins

amino acids 475-504

FIGURE 205

TGCCTGGCCTGCCTTGTCACAATGCCGCTTACTCTGCTTCCAGGTTGCCCTGCCTTG CAGA
GGAAANCNTCGGGACTACACCNCAAGTGACATGAACCTGCTGCTGGCCGTCTTCCTGCTG
GACACGAGCTTCCTGCTCAGCGNAGCCGGTGGCCCTGACAGGCTCTGAAGGCTGGCTGCCGA
GCCAGTGCCATCTTCCTGCACTTCTCCTGCTCACCTGCCTTTCTGGATGGGCCTCGAGGGG
TACAACTCTACCGACTCGTGGTGGAGGTCTTTGGCACCTATGTCCCTGGCTACCTACTCAA
GCTGAGCGCCATGGGCTGGGGCTTCCCATCTTTCTGGTGACGCTGGTGGCCCTGGTGGATG
TGGACAACCTATGGCCCATCATCTTGGCTGTGCATAGGACTCCAGAGGGCGTCATCTACCCCT
TCCATGTGCTGGATCCGGGACTCCCTGGTCAGCTACATCACCAACCTGGGCCTCTTCAGCCT
GGTGTTCCTGTTCAACATGG

0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
1.1
1.2
1.3
1.4
1.5
1.6
1.7
1.8
1.9
2.0

FIGURE 206

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCTGGTTCAGGTCCAGGTTTTCCTTTGA
TCCTTTTCAAAAACCTGGAGACACAGAAAGGGCTCTAGGAAAAAGTTTGGATGGGATTTATGTGGAAACCTACCCCT
GCGATTCTCTGCTGCCAGAGCAGGCTCGGCGCTTCCACCCAGTGCAGCCTTCCCTCTGGCGGTGGTGAAGAGAGAC
TCGGGAGTCGCTGCTCCAAAGTGCCCGCGGTGAGTGAGCTCTCACCCAGTCAAGCCAAATGAGCGCTCTTCGGGGC
TTCTCTGCTGACATCTGCCCTGGCGCGCCAGAGACAGGGGACTCAGGCGGAATCCAACCTGGATGTAAATTC
AGTTTTCCAGCAACAAGGAACAGAACGGAGTACAAGATCTCAGCATGAGAGAATTTATCTGTGTCTACTAATG
GAAGTATTCACAGCCCAAGGTTTTCTCATACTTATCCAAGAAATACGGTCTTGGTATGGAGATTAGTAGCAGTAG
AGGAAATGTATGGATACAACCTACGTTTGTATGAAGATTGGGCTTGAAGACCCAGAAGATGACATATGCCAATG
ATGATTTTGTAGAAGTTGAGGAACCCAGTGTATGGAACCTATATTAGGGCGCTGGTGTGGTTCTGTGATCTGTACCAG
GAAACAGATTTCTAAAGGAAATCAATTAGGATAAAGATTGTATCTGATGAATATTTCTCTCTGAACCAGGCT
TCTGCATCCACTACAACATGTGTCATGCCACAATTACACAGAAGCTGTGAGTCCCTTCAGTGCTACCCCTTCAGCTT
TGCCACTGGACCTGCTTAATAATGCTATAACTGCCCTTAGTACTCTTGAAGACCTTATTCGATATCTTGAACCGAG
AGAGATGGCAGTTGGACTTAGAAGATCTATATAGGCCAACTTGGCAACTCTTGGCAAGCGCTTTTGGTGTGGAA
GAAATCCAGAGTGGTGGATCTGAACCTTCTAACACAGAGAGGTAAAGATTATACAGCTGCACACCTCGTAACCTTCT
CAGTGTCCATAAGGGAAGAACCTAAAGAGAACCGATACCATTTCTGGCCAGGTTGTCTCTGGTAAACGCTGTG
GTGGGAACTGTGCCCTGTGTCTCCACAATTGCAATGAATGTCAATGTGTCCCAAGCAAAAGTTACTAAAAAATACC
ACGAGTCCCTTCAGTTGAGACCAAGAACCGGTGTGAGGGGATGACAAATCACTCACGACGTGGCGCTGGAGC
ACCATGAGGAGTGTGACTGTGTGTGCAGAGGGAGCACAGGAGGATAGCCGCATCACCCAGCAGCTCTTGGCCCA
GAGCTGTGCAGTGCAGTGGCTGATTCTTATAGAGAACGTATGCGTTATCTCCATCTCTTAATCTCAGTTGTTTGCT
TCAAGGACCTTTCATCTTCAGGATTTCAGTGCACTTCTGAAAGAGGAGACATCAACAGAATTAGGAGTTGTGCA
ACAGCTCTTTTGGAGGAGGCCATAAGAGACAGGAGAAAAGGCTTCAATCGTGGAAAGAAAATTAATATGTGTAT
TAATAGATCACAGCTAGTTTTCAGAGTTACCATGTACGTATTCCAAGTGGGTTCTGTATTTTCAGTTCTTTTC
GATACGGCTTAGGGTAATGTCACTACAGGAAAAAACTGTGCAAGTGAACCTGATTCGGTTCGCTTGCCTTAAC
TCTAAAGCTCCATGTCTCGGGCTTAAATCGTATAAAATCTGGATTTTTTTTTTTTTTTTGTGCTCATATTCACAT
ATGTAAACCAAGACATCTCTATGTACTACAAACCTGGTTTTTAAAGGAACATATGTTGCTATGAATTAACCTGT
GTCACTGTGATAGGACAGACTGGATTTTTCATATTTCTTATTAATAATTTCTGCCATTAGAGAAGAGAACTACA
TTCATGGTTTGAAGAGATAAACCTGAAAAGAAAGAGTGGCCTTATCTTCACTTATCGATAAGTCAGTTATTATTG
TTTCATTGTGACATTTTATATTTCTCTTTTGACATTATAACTGTGGCTTTTCTAATCTGTGTAATATATCT
ATTTTTACCAAGGATTTTAAATATCTTTTTTATGACAACCTAGATCAACTATTTTTAGCTTTGGTAAATTTTTCT
AAACAATTTGTTATAGCCAGAGGAACAAGATGATATAAAATATGTGTGCTGTGCAAAAATACATGTATTTCA
TTCCTGATGGTGTAGAGTTAGATTAATCTGCATTTTAAAAAACTGAATTTGAATAGAAATTTGGTAAAGTTGCAAA
GACTTTTTGAAAAATAATTAATATCATATCTTCCATTCTGTATTGGAGATGAAAATAAAAAGCAACTATATGA
AAGTAGACATTCAGATCCAGCATTACTAACCATTCTCTTTTGGGAAATCTGAGCTTAGCTCAGAAAAAATCT
AAAGACCTTGAAAAAGACTTGGCGACTTCCGTATAAGCGTGTGCTGTGCTGACATAGGACACATCTCATTTTA
TTGTGATGTGTGGTTTTTATTATCTTAACTCTGTTCATACACTTTGTATAAATACATGATATTTTTATGTACA
GAAGTATGTCTCTTAAACGATCTATTGTACTCTGCGCAATTTAAAAAGAAATCAGTAAATATTGTCTTGT
AAAATGCTTAAATATNGTGCCTAGGTATATGTGTGACTATTGAACTCAAAATGTATTGAATCATCAAAATAAGA
ATGTGGCTATTTTGGGAGAAAAATTAAAAAAAAAAAAAAAAAAAGTTTAGGGATACAGGGTAAATGCGGC

FIGURE 207

MSLFGLLLLSALAGQRQGTQAESNLSSKFQFSSNKEQNGVQDPQHERIITVSTNGSIHSPR
FPHTYPRNTVLVWRLVAVEENVVIQLTFDERFGLEDPEDDICKYDFVEVEEPSDGTILGRWC
GSGTVPGKQISKGNQIRIRFVSDEYFPSEPGFCIHYNI VMPQFTEAVSPSVLPSPALPLDLL
NNAITAFSTLEDLIRYLEPERWQLDLEDLYRPTWQLLGKAFVFGKRSRVVDLNLLEEVRLY
SCTPRNFSVSIREELKRTDTIFWPGCLLVKRCGGNCACCLHNCNECQCVP SKVTKKYHEVLQ
LRPKTGVRGLHKSLTDVALEHHEECDVCVRGSTGG

Signal sequence:

amino acids 1-14

100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120

FIGURE 208

CCCATCTCAAGCTGATCTTGGCACCTCTCATGCTCTGCTCTCTTCAACCAGACCTCTACATTCATTTTGGAAGA
 AGACTAAAAATGGTGTTCCTCAATGTGGACACTGAAGAGACAAATCTTATCCITTTTAAACATAATCTCAATTTCC
 AAACCTCTTGGGGCTAGATGGTTTCTTAAACCTCTGCCCTGTGATGTCACTCTGGATGTCTCAAAGAACCATGTG
 ATCGTGGACTGCACAGACAAGCAITTTGACAGAAATTCCTGGAGGTATTCACAGAACACACAGAACCTCACCCCTC
 ACCATTAAACACATACAGACATCTCCCCAGCGTCTTTACAGACTGGACCATCTGGTAGAGATCGATTTCAGA
 TGCAACTGTGTACCTATTCCATCTGGGGTCAAAAAACAACATGTGCATCAAGAGGCTGCAGATTAAACCAGAAAG
 TTTAGTGGACTCACTTATTTTAAATCCCTTTACCTGGATGGAAACAGCTACTAGAGATACCGCAGGGCTCCGC
 CCTAGCTTACAGCTTCTCAGCCTTGAGGCCAACCAACATCTTTTCCATCAGAAAAGAGAATTAAACAGAACTGGCC
 AACATGAAAAATCTCTACCTGGGCCAAAACTGTATTATCGAAATCCTGTGTATTGTTTATATTCAATAGAGAAA
 GAGTCCTCTTCAATCTGACAAAGTTAAAGTGCTCTCCCTGAAAGATAACAAATGTGCACGCCGTCCCTCTCTGTT
 TTGCGCTCTACTTTAAACAGAACTATATCTCTCAACAACATGATTGCAAAATCCAGAGAGATGATTTTAAATAC
 CTCAACCAATTACAAATTTCTGACCTAAGTGGAAATGGCCCTCGTTGTTATAATGSCCATTTCCTTGTGCGCGG
 TGTAATAAATAATCTCCCCCTACAGATCCCTGTAAATGCTTTTGTATGCGCTGACAGAAATAAAGTTTTCAGCTTA
 CACAGTAACCTCTCTCAGCATCTGCCCCCAAGATGGTTTAAAGAACATCAACAACTCCAGGAACCTGGATCTGTCC
 CAAAACCTCTTGCCCAAAGAAATTTGGGGATGCTAAATTTCTGCAATTTCTCCCGAGCTCATCCAATTTGGATCTG
 TCTTTCAATTTTGAACCTTCAGGTCTATCGTGCATCTATGAATCTATCAACAGCAATTTCTTCACTGAAAAGCGTG
 AAAATCTCGCGATCAGAGGATATGTCTTTAAAGAGTTGAAAAGCTTTAAACCTCTCGCCATTACATATACTCTCAA
 AATCTTGAAGTTCTTGATCTTGGCACTAACTTTATAAAAAATTGCTAACCTCAGCATGTTTAAACAATTTTAAAGA
 CTGAAAGCTCATAGATCTTTTCAGTGAATAAAATATCACCTTCAGGAGATTCAAGTGAAGTTGGCTTTCTGCTCAAA
 GCGCAAACTCTGTAGAAAAGTTATGAACCCAGGTCTTGGAAACAATTACATTATTCAGATATGATAAGTATGCA
 AGGAGTTGCAGATTCAAAACAAAGAGGCTTTCTTCAATGCTGTAAATGAAGCTGCTACAAGTATGGGCAGACC
 TTGGATCTAAGTAAAAATAGTATATTTTGTCAAGTCCTCTGATTTTCAGCATCTTCTCTCTCAAAATGGCTG
 AATCTGTCAAGAAATCTCATTAGCCAAACTCTTAATGGCAGTGAATTCACACCTTTAGCAGAGCTGAGATATTG
 GACTTCTCCAACAACCGGCTTGATTTACTCCATTCAACAGCAATTTGAAGAGCTTCACAAAGTGAAGTTCTGGAT
 ATAAGCAGTAATAGCCATTATTTTCAATCAGAAGGAATTAATCATATGCTAAACTTTACCAAGAACCTAAAGGTT
 CTGCAGAACTGATGATGAACGACAAATGACATCTCTCTCCCAACAGCAGGACCATGGAGAGTGAGTCTCTTAGA
 ACTCTGGAATTCAGAGGAATCACTTAGATGTTTATGGAGAGAAGGTGATAACAGATATTCAAAATATTCAAG
 AATCTGCTAAAAATTAGAGGAATTAGACATCTCTAAAAATTCCTAAGTTTCTTGGCTCTGGAGTTTGTGATGGT
 ATGCCCTCCAAATCTAAAGAATCTCTCTTTGGCCAAAAATGGGCTCAAACTCTTCAGTTGGAAGAACTCCAGTGT
 CTAAGAAACCTGGAACTTTGGACCTCAGCCAAACCAACTGACCACTGCTCCCTGAGAGATTATCCAACCTGTTCC
 AGAAGCTCAAGAATCTGATCTCTTAAGAATAATCAAACTCAGGAGCTCTGACGAAGTATTTTCTACAGATGCTCTC
 CAGTTGCGATATCTGATCTCAGCTCAAAATAAAATCCAGATGATCCAAAGAGCCAGCTTCCCAAGAAATGTCCCTC
 AACAACTGGAAGATGTGTCTTTGCATCATAATCGGTTTCTGTGCACCTGTGATGCTGTGTGTTTGTCTGTGTTG
 GTTAACCATACGGAGGTCTTCTTACCTGGCCACAGATGTGACTTGTGTGGGGGCCAGGACACACAAGGGC
 CAAAGTGTGATCTCCCTGGATCTGTACACCTGTGAGTTAGATCTGACTAACTGATTCTGTGTTCTCACTTTCTCAT
 TCTGTATCTCTCTTTCTCATGGTGATGATGACAGCAAGTCACTCTTATTTCTGGATGTGTGGGTATATTACCAT
 TTTCTGTAAAGCCCAAGTAAAGGGATTCAGCGTCTTAATATCACAGAGCTGTGCTATGATGCTTTTATTTGTGAT
 GACACTAAGGCCAGCTGTGACCGAGTGGGTTTTTGGCTGAGCTGGGCTTGAAGTCTGGAAGCCCAAGAGAGAA
 CATTTTAAATTTATGCTCAGGAGAAAGGACTGGTTACCAAGGGCAGCCAGTTCTGGAAGAACTTTCCAGAGACATA
 CAGCTTGAAGCAAAAGCAGTGTGTTGTGATCAGACAAAGTCTCAAGAGCTGAAATTTTAAAGTAGCATTTTAC
 TTGTCCTCATCAGAGGCTCATGTGATGAAAAGTTGATGTGATTTATCTGTATTTCTTGAAGAGCCCTTTCAAGAG
 TCCAGGTTCTCTCAGCTCCGGAAGAGGCTCTGTGGAGTTCTGTCTTGTGAGTGGCCCAACAAACCCGCAAGCTCAC
 CCTACTCTCTGGCAGTGTCTAAAGACGSCCTGGCCACAGACATCATGTGGCTCATGTCCAGTGTCTCAAGGAA
 ACGGCTCTAGCCCTTCTTTGCAAAACACACTGCGCTAGTTTACCAAGGAGAGGCGCTGGC

FIGURE 209

MVFPMTLKRQILILFNIILISKLLGARWFPKTLPCDVTLDVPHVIVDCTDKHLTEIPGG
IPTNTNTLTLTINHIPDISPASFHRLDHLVEIDFRNCVPIPLGSKNNMCIKRLQIKPRSFS
GLTYLKSLYLDGNQLLEIPQGLPPSLQLLSLEANNIFSIRKENLTELANIEILYLGQNCYYR
NPCYVSYSIEKDAFLNLTCLKVLSLKDNNVTAVPTVLPSTLTELYLYNNMIAKIQEDDFNNL
NQLQILDLSGNCPRCYNAPFPCAPCKNNSPLQIPVNAFDALTELKVLRLHSNSLQHVPPRWF
KNINKLQELDSLQNFLAKEIGDAKFLHFLPSLIQLDLSFNELQVYRASMNLSQAFSSLKSL
KILIRIGYVFKELKSFNLSPLHNLQNLEVLDLGTNFIKIANLSMFKQFKRLKVIDLSVNKIS
PSGDSSEVGFCSNARTSVESYEPQVLEQLHYFRYDKYARSCRFKNKEASFMSVNESCYKYGQ
TLDLSKNSIFFVKSSDFQHLSFLKCLNLSGNLISQTLNGSEFQPLAELRYLDFSNNRLDLLH
STAFEELHKLEVLDISSNSHYFQSEGITHMLNFTKNLKVQLKLMMDNDIISSTSRTMESES
LRTLEFRGNHLDVLWREGDNRYLQLFKNLLKLEELDISKNSLSFLPSGVFDGMPNKLKNSL
AKNGLKSFSWKKLQCLKNLETLDLSHNQLTTVPERLSNCSRSLKNLILKNNQIRSLTKYFLQ
DAFQLRYLDLSSNKIQMIQKTSFPENVLNNLKMLLLHHNRFLCTCDAVWFVWVNHTVETIP
YLATDVTCVGPAGAHKQSVISLDLYTCELDLTNLILFSLISVSLFLMVMMTASHLYFWDVW
YIYHFCKAKIKGYQRLISPDCCYDAFIVYDTKDPVTEWVLAELVAKLEDPREKHFNLCLEE
RDWLPGQPVLLENLSQSIQLSKKTVFVMTDKYAKTENFKIAFYLSHQRLMDEKVDVILIFLE
KPFQKSKFLQLRKRLCGSSVLEWPTNPQAHYPYFWQCLKNALATDNHVAYSQVFKETV

Signal sequence:

amino acids 1-26

Transmembrane domain:

amino acids 840-860

FIGURE 210

GGGTACCATTCTGCGCTGCTGCAAGTTACGGAATGAAAAATTAGAACACAGAAAAATGAGAAAAACATGTTCTCTTC
AGTCGTCATGCTGACCTGCAATTTTCTGCTAATATCTGGTCTCTGTGAGTTATGCGCCGGAAGAAAAATTTTCTA
GAAGCTATCCTTTGTGATGAGAAAAAGCAAAATGACTCAGTTATTGCAGAGTGCAGCAATCGTGCAGCTACGGAAG
TTCCCCAAGCCGGTGGCAAAATATGTGACAGAAGCTAGACCTGTCTGATAATTTTCATCACACACATAACGAATGAAT
CATTTCAAGGCGTGCRAAAATCTCACTAAAAATAAATCTAAACACACACCCCAATGTGACAGCACACAGAACGGAATC
CCGGTATACAATCAATAGGCTTGAATATCACAGACGGGGCAATTCCTCAACCTAAAAAACCTTAAGGGAGTTACTGCT
TTGAAGACACACCAAGTTACCCCAATACCCCTCTGGTTTGGCCAGAGTCTTTGACAGAACTTAGTCTTAATTCAAAAACA
ATATATACAAACATAACTTAAGAGGGGCATTTCAAGACTTATAAACTTGAAGAAATCTTAATTTGGCCGTGGAACCTGCT
ATTTTAACAAAGTTTGGCAGAAAACTAAACATAGAAGATGGAGTATTTGAACCGCTGACAAATTTGGAGTGTCTAT
CACTATCTTTCAATTTCTTGTTCACACAGTGCACCCCAAACCTGCCAAGCTCCCTACGCAAACTTTTCTGAGCAACA
CCAGATCAAAATCAATCTGTGGAAGAAGATTCTCAAGGGATTGATAAATTTAACTATAGATTATAAGCGGAACT
GTCCGAGGTGCTCAATGCGCCCAATTTCCATGCGTGCCCTGTGATGGTGGTCTCAATTAATATAGATCGTTTTG
CTTTCAAACTTGCACCAACTTGATACTTAAACCTCTAGCACTCCCTCAGGAAGAAATTAATGCTGCCCTGGT
TTAAAAATATGCTCACTGAAAGGTGCTGGATCTTGAATTCAACTATTAGTGGGAGAAATAGTCTCTGGGCAAT
TTTTAACGATGCTGCCCGCTAGAAATACTTGACTTGTCTTTAACTATATAAGGGAGATATCCACAGCAAT
TTAATATTTCAGAAACTTCTCTAAACCTTTTGTCTCTACGGGCATTTGCATTTAAGAGTTATGTGTTCCAGGAAC
TCAGAGAAGATGATTTCCAGCCCTGATGCAGCTTCCAAACTTATCGACTATCAACTTGGGTATTAATTTATTATTA
AGCAAACTCGATTTCAAACCTTTCAAATAATTTCTCAACTCTGGAATTAATTTACTGTGCAGAAACAGAAATATCAC
CGTTGGTAAAGATACCCGGCAGAGTTATGCAAAATAGTTCTCTTTCAACGTCTATTCGGGAAACGAGCGCTCAA
CAAGTTTGGATTTGACCCCAATTCGAACCTTTTATCATTTCAACCGCTCTTTAATAAGCCACAAATGTGCTGCTTT
ATGAAAAAGCCTTAGATTTAAGCTCAACAGATATTTCTTCATTGGGCAAAACCAATTTGAATAATCTTCTTGACA
TGCCCTGTTTAAATCTGTCTGCAAAATAGCAATGCTCAAGTGTAAAGTGAAGTGAATTTTCAAGCAATCTCTCATG
TCAAAATATTTGGATTTGACAAACAAATAGACTAGACTTTGATAATGCTAGTGCTCTTACTGAATTTGTCGAGCTGG
AAGTCTAGATCTCAGCTATAATTCACACTATTTCAAGATAGCAGGCGTAACACATCATCTAGAAATTTATTCAA
ACTTCCAAAGCTTAAAGTTTAAACTTGAGCCACAACAACATTTATCTTTACCAAGATAAGTATAACCTGGAA
CGAAGTCCCGTGTAGAAATAGTTTTCAGTGGCAATCGCCTTGACATTTTGTGGAATGATGACACACAGGTATA
TCTCCATTTTCAAGAGTCTCAGAAATCTGACACGCTTGGATTTATCCCTTAAAGGCTGAAGCACATCCCAAATG
AAGCATCTCTAATTTGGCCAGCGAGTCTCACTGAACTACATATAAATGATAAATGTTAAAGTTTTTAACTGGA
CATTTACTCCAGCAGTTTCTCGTCTCGAGTTGCTTGACTTACGTGGAACAAACTACTCTTTTAACTGATAGCC
TATCTGACTTTACATCTTCCCTTCGGACACTGCTGCTGAGTCAATAACAGGATTTCCCACTACCCCTCGGCTTTC
TTTCTGAAGTCAGTAGTCTGAAGCACCCTCGATTTAAGTTCCAATCTGTAACAAACAAATCAACAAATCCGCACTG
AAACTAAGACCAACCAAAATTTCTATGTTGGAACCTACAACGAAACCCCTTTGAATGCACCTGTGACATTTGGAG
ATTTCCGAAGATGGATGGATGAACATCTGAATGTCAAATTTCCAGACTGGTATGATGTCTTGTGCCAGCTCTG
GGGATCAAAGGGGAAGAGTATGTGAGTCTGGAGCTTAACAACTTGTGTTTTCAGATGTCACGTGAGCTGATTTAT
TTTTCTTCACTGTTCTTTATCACCACCAATGGTTATGTTGGCTGCCCTGGCTTCAACATTTTGTTTTTCGGAATGTT
GGTTTATATAATGTGTGTTTGAAGTAAAGGCTAAGGCTACAGTCTTTTCCACATCCCAAATTTTCTATGATG
CTTACATTTTCTATGACACCAAGAGTGCCTCTGTTTACTGACTGGGTGATAAATAGACTGCGCTTCAACCTTGAAG
AGAGCCGAGACAAAGCTTCTCCCTTGTCTAGAGGAGAGGATTTGGGACCGGATTTGGCCATCATCGACAAC
TCATCGAGAGCTCAACCAAGACGAAGAAACAGTATTTGTTTAAACAAAAAATATGCAAAAGCTGGAATCTTAA
AAACAGCTTTTACTTTGGCTTTGACAGGCTAATGGATGAGAACTGAATGATTAATTTATCTCTCTGGAGC
CAGTGTACAGCATCTCAGTATTTAGGCTACCGCAGCGGATCTGTAAGAGCTCCATCTCTCAGTGGCTGCACA
ACCCGAAGACAGAGGCTGTTTGTGGCAAACTCTGAGAAATGTGCTTGTGACTGGAATAATGATTCAGGCTATAACA
ATATGATGTGCTGATCCATTAAGCAATACTAACTGACGTTAAGTCATGATTTCCGCGCATATAAAGAGTCGAAG
GAATGACATCTGATATTAGTTATCTATTTGCTATGTAAACAAATATCCCAAACTTAGTGGTTTAAACACACACA
TTTGTCTGCCCCAGCTTTTGTAGGCTCAGGAGTCCAGGCCAGCATAACTGGGCTCTCTGTCAGGCTCTCAG
AGGCTGCAATGTAGTGTCTTCCACAGAGACATAGGCATCACTGGGCTCACACTCATGTGGTGTGTTTCTTGGAATCA
ATTTCTCTGGGCTATTTGGCCAGGCTATCATCTCATGTAAGCCATGCGAGCTCTCCCAACAGGCGAGTCTGCTC
ATCAGAGCTAGCAAAAGAGAGAGTGTGTAGCAAGATGAAGTCAAACTTTTGTAACTGAAATCAAAAGATGAT
ATCTCATCATCTTTGGCCATTTTATTTGTAGAAGTAAACACAGGCTCCACACAGCTCATGGGAGTGCAC
TCAGTCTGAGGAAAAACAGCTGAAGACCAAGATGTTGAGCTCTGATGCTTCAGTTGTCATCAACTATTTTCCCT
TGACTGTGCTCTGGGATGGCTCTCATCTTGATGATAGATTGTAATATCAGGAGGCGAGGACTCACTGTGGAC
ATCTTAGCAGTGACTCAACACATCTTTTCAATATCAAGAAATTTTGGCACTGACTATGATGGTCTGATGATG
TTAAGCTGTGTTGTTTATTTTATCATATACTATGGCTACATGGTTATTAATGCTGTGGTTGCTGTGGTTTAT
TTACAGTGTGCTTTTACAAATATTTGCTGTAACATTTGACTTCAAGGTTTAGATGCCATTTAAGAACTGAGATGG
ATAGCTTTTAAAGCATCTTTTACTCTTCAACATTTTAAAGTATGCAAGCTAAATTCAGGCTTTTGGCTCTATA
TTGTTAAATGCCATTTGCTGTAATCTTAAATGAATGAATAAAATGTTTCATTTTCAAAAAA

FIGURE 211

MENMFLOSSMLTCIFLLISGSCELCAEENFSRSYPCDEKKQNDSVIAECSNRRLQEVQPQTVG
KYVTELDLSDNFITHITNESFQGLQNLTKINLNHNPNVQHQNNGPIQSNGLNITDGAFLNL
KNLRELLLEDNQLPQIPSGLPESLTELSLIQNNIYNITKEGISRLINLKNLYLAWNCYFNKV
CEKTNIEDGVFETLTNLLELLSLSFNSLSHVPPKLPSSLRKFLSNTQIKYISEDFKGLINL
TLDLDSGNCPRCFNAPFPCVPCDGGASINIDRFQNLTLQRLYNLSSTSLRKINAAWFKNM
PHLKVLDFENYLVGEIVSGAFLTMLPRLEILDLSFNYYIKGSYPQHINISRNFSKLLSLRAL
HLRGYVFQELREDDFQPLMQLPNLSTINLGINFIKQIDFKLFQNFSLNLEIYYLSENRI SPLV
KDTRQSYANSSSFQRHIRKRSTDFEFDPHSNFIYHFRPLIKPQCAAYGKALDLSLNSIFFI
GPNQFENLPDIACNLNSANSNAQVLSGTEFSAPHVKYLDLTNNRLDFDNASALTELDLEV
LDLSYNSHYFRIAGVTHHLEFIQNFNLKVLNLSHNNIYTLTDKYNLESKSLVELVFSGNRL
DILWNDDDNRYISIFKGLKNLTRLDSLNLRLKHIPNEAFLNLPASTEHLHINDNMLKFFNWT
LLQQFPRLELLDLRGNKLLFLTDSLSDFTSSLRTLSSHNRISHLPSGFLSEVSSLKHLDSL
SNLLKTINKSALETKTTTKLSMLELHGNPFECTCDIGDFRRWMDEHLNVKIPRLVDVICASP
GDQRGKSIVSLELTTCVSDVTAVILFFFTFFITMVMLAALAHHLFYWDVWFIYNVCLAKVK
GYRSLSTSQTIFYDAYISYDTKDASVTDWVINELRHYLEESRDKNVLLCLEERDWDPLAIIID
NLMQSINQSKKTVFVLTKKYAKSWNFKTA FYLALQRLMDENMDVIIIFILEPVLQHSQYLRL
RQRICKSSILQWPDNPKAEGLFWQTLRNVVLTTENDSRYNMYVDSIKQY

Signal sequence:

amino acids 1-26

Transmembrane domain:

amino acids 826-848

FIGURE 212

CCAGGTCCAACCTGCACCTCGGTTCTATCGATTGAATCCCCGGGGATCCTCTAGAGATCCCT
CGACCTCGACCCACGCGTCCGCCAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGA
CAGGCCAGGCAGGTGGGCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCTGAGGCCCCAGC
AAGGGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATTGGCCACGCTGGGC
TCCAGCAGCATCAGCAGCCCCAGGACCGGGAGGCACAGGTGGCCCCCACCACCCGGAGGA
GCAGCTCCTGCCCTGTCCGGGGGATGACTGATTCTCTCTCCGCCAGGCCACCCAGAGGAGAA
GGCCACCCCGCTGGAGGCACAGGCCATGAGGGGCTCTCAGGAGGTGCTGCTGATGTGGCTT
CTGGTGTGGCAGTGGGCGGCACAGAGCACGCTTACCGCCCCGGCGTAGGGTGTGTGCTGT
CCGGGCTCACGGGACCTGTCTCCGAGTCGTCGTGCAGCGTGTGTACCAGCCCTTCTCTCA
CCACCTGCAGCGGCACCGGGCTGCAGCACCTACCGAACCATCTATAGGACCGCTTACCGC
CGCAGCCCTGGGCTGGCCCCCTGCCAGGCCTCGCTACGCGTGTGCCCCGGCTGGAAGAGGAC
CAGCGGGCTTCTCGGGGCTGTGGAGCAGCAATATGCCAGCGCCATGCCGGAACGGAGGGA
GCTGTGTCCAGCCTGGCCGCTGCCGCTGCCCTGCAGGATGGCGGGGTGACACTTGCCAGTCA
GATGTGGATGAATGCAGTGTCTAGGAGGGGCGGCTGTCCCAGCGCTGCATCAACACCGCCGG
CAGTTACTGGTGCCAGTGTGGGAGGGGCACAGCCTGTCTGCAGACGGTACACTCTGTGTGC
CCAAGGGAGGGCCCCCAGGGTGGCCCCAACCCGACAGGAGTGGACAGTGAATGAAGGAA
GAAGTGCAGAGGCTGCAGTCCAGGTGGACCTGTGGAGGAGAAGTGCAGCTGGTGTGTGGC
CCCACTGCACAGCCTGGCCTGCAGGCACTGGAGCATGGGCTCCCGAACCCGGCAGCCTCC
TGGTGCACCTCCTTCCAGCAGCTCGGCCGCATCGACTCCCTGAGCAGCAGATTTCCTTCTCTG
GAGGAGCAGCTGGGGTCTGTCTCTGCAAGAAAGACTCGTGACTGCCCAGCGCCCCAGGCTG
GACTGAGCCCCCTACGCGCCCTGCAGCCCCCATGCCCTGCCCAACATGTGGGGGTCCAG
AAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGCAGGGCCTTCTCTCTTCTCTCTCCCC
TTCCTCGGGAGGCTCCCCAGACCCTGGCATGGGATGGGCTGGGATCTTCTCTGTGAATCCAC
CCCTGGCTACCCCCACCTGGCTACCCCAACGGCATCCCAAGGCCAGGTGGGCCCTCAGCTG
AGGGAAGGTACGAGCTCCCTGCTGGAGCCTGGGACCCATGGCACAGGCCAGGCAGCCCGGAG
GCTGGGTGGGGCTCAGTGGGGGCTGTGCTGACCCCCAGCACAAATAAAATGAAACGTGA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGCACTCTAGAGT
CGACCTGCAGAAGCTTGGCCGCATGGCCCACTTGTTTATTGCAGCTTATAATGGTTACAAAT

FIGURE 213

MRGSQEVLLMWLLVLAVGGTEHAYRPGRRVCAVRAHGDPVSESFVQRVYQPFLTTCDGHRAC
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICPPPCRNGGSCVQPGRCR
CPAGWRGDTQCSDVDECSARRGGCPQRCINTAGSYWCQWEGHSLSADGTL CVPKGGPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEKQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

1-19

FIGURE 214

GCCAGGCAGGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGCAAG
GGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGCTCC
AGCAGCATCAGAGCAGCCCCGTGTGGTTGGCAGCAAAGTTCAGCTTGGCTGGGCCCGCTGTGA
GGGGCTTCGCGCTACGCCCTGCGGTGTCCCGAGGGCTGAGGTCTCCTCATCTTCTCCCTAGC
AGTGGATGAGCAACCCACGCGGGCCCCGGGGAGGGGAAC TGGCCCCGAGGGAGAGGAACCCC
AAAGCCACATCTGTAGCCAGGATGAGCAGTGTGAATCCAGGCAGCCCCCAGGACCGGGGAGG
CACAGGTGGCCCCCACCACCCGAGGAGCAGCTCCTGCCCTGTCCGGGGGATGACTGATTC
TCCTCCGCCAGGCCACCCAGAGGAGAAGGCCACCCGCCCTGGAGGCACAGGCCATGAGGGGC
TCTCAGGAGGTGCTGTGTGTGTGGCTTCTGGTGTGGCAGTGGGCGGCACAGAGCACGCCTA
CCGCCCCGGCCGTAGGTTGTGTGTGTCCGGGCTCACGGGGACCTGTCTCCGAGTCGTTCTG
TGCAGCGTGTGTACCAGCCCTTCCTCACCACCTGCGACGGGCACCGGGCTGCAGCACCTAC
CGAACCATCTATAGGACCGCTACCGCGCAGCCCTGGGCTGGCCCCCTGCCAGGCCTCGTA
CGCGTGTGCCCCCGCTGGAAGAGGACCAGCGGGCTTCCTGGGGCCTGTGGAGCAGCAATAT
GCCAGCCGCATGCCGGAACGGAGGGAGCTGTGTCCAGCCTGGCCGCTGCCGCTGCCCTGCA
GGATGGCGGGGTGACACTTGCCAGTCAGATGTGGATGAATGCAGTGTCTAGGAGGGGCGGCTG
TCCCCAGCGCTGCATCAACACCGCCGCGAGTTACTGGTGCAGTGTGGGAGGGGCACAGCC
TGTCTGCAGACGGTACACTCTGTGTGCCCAAGGAGGGCCCCCAGGGGTGGCCCCCAACCCG
ACAGGAGTGGACAGTGCAATGAAGGAAGAAGTGCAGAGGTGTCAGTCCAGGGTGGACCTGCT
GGAGGAGAAGCTGCAGCTGGTGTGGCCCCACTGCACAGCTTGGCCTCGCAGGCACTGGAGC
ATGGGCTCCCGGACCCCGGCAGCCTCCTGGTGCACCTCCTTCCAGCAGTCCGGCCGCATCGAC
TCCTTGAGCGAGCAGATTTCTTCTTCTGGAGGAGCAGCTGGGGTCTCTCTCTGCAAGAAAGA
CTCGTGACTGCCCAGCGCTCCAGGCTGGACTGAGCCCCCTACGCCGCCCTGCAGCCCCCATG
CCCTGCCCAACATGTGCGGGGTCCAGAAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGC
AGGGCTTCCTCCTCTTCTCCTCCCTTCTCTCGGAGGCTCCCAGACCTGGCATGGGAT
GGGCTGGGATCTTCTCTGTGAATCCACCCCTGGCTACCCCAACCTGGCTACCCCAACGGCA
TCCCAAGGCCAGGTGGACCTCAGCTGAGGGAAGGTACGAGCTCCCTGCTGGAGCTGGGAC
CCATGGCACAGGCCAGGCAGCCCGGAGGCTGGGTGGGGCCTCAGTGGGGGTGTGCTGCCTGAC
CCCCAGCACAAATAAAATGAAACGTG

Figure 1

Signal sequence:

1-19

Signal sequence:

1-19

FIGURE 218

GGTTGCCACAGCTGGTTTAGGGCCCCGACCATTGGGGCCCCCTTGTCTCAGGAGGAGACAGCCTCCCGGCCCGGGGAG
GACAAGTCGCTGCCACCTTTGGCTGCCGACGTGATTCCTTGGGACGGTCCGTTTCTCTCCGCTCAGCTGCCGCGCG
AGTTTGGGTCTCCGCTGTTTTCAGGCCCGCTCCCCCTTCCTGGTCTCCCTTCTCCGCTGGCCCGGTTTATCGGGAGG
AGATTGTCTTCCAGGCTAGCAATTGGACTTTTGTATGATGTTTGACCCAGCGCGCAGGAATAGCAGGCAACGTGAT
TTCAAAGCTGGGCTCAGCCTCTGTTTCTCTCTCGTGAATCGCAAACCCATTTTGGAGCAGGAATCCAATCA
TGTCTCTGTATGGTGGTGAGAAAGAGGTGACACGGAAATGGGAGAAATCCCAGGCGAGGAACACCTTTTGTCTGTG
ATGGCCGCGCTCATGATGGCCCGGCAAAGGGCATTTCCTACCTGACCCCTTTTCTCATCTCTGGGGACATGTACAC
TCCTCTCTGCCCTTTGAGTGCCTCTACCTGGCTGTTTCTCAGCTCTCTCTGCCATCCCTGTATTGTGCTGCCATGTCTCT
TCCTTTTCTCATTGGCTACACTGTTGAGGACCAAGCTTCACTGACCTGGAGTGATTCTCTCGGCGCTACCCAGATG
AAGCAGCTTTCATAGAAATGGAGATAGAAAGCTACCAATGTGTCGGTGGCCGAGCGGACCGGACCAACGCTCTCGTA
TCAAGAAATTTCCAGATAAAACAACCAAGATGTGAAACTGMAATACGTGTACACATGCAAGATCTTCCGGCTCCCC
GGGCTCCCATTTGCAGCATCTGTGACAACTGTGTGGAGCGCTTCGACCATCACTGCCCCCTGGGTGGGGAATGTG
TTGGAAAGAGGAACACCGCTACTTCTACCTCTTCATCCTTTCTCTCCCTCCTCACAATCTATGTCTTCGCT
TCAACATCGCTATGTGGCCCTCAAATCTTTGAAATTTGGCTCTTGGAGACATTGAAAGAAACTCCTGGAACTG
TTCTAGAAGTCTCTATTGCTCTTTTCACTCTGTGTCGTCGTGGGAGTACTGGATTTCATACTTCTCTCGTGT
CTCTCAACAGACACCACTGAACTCAAATGAGATGCCGAGGACAGCAGCACTCCCGAAGAGATGCCACCTCCAG
ATGGCAATATTTGTAAGAACTGCTGTGAAGTCTGTGTGCCCTTTGCCCCCGAGTGTGCTGGATCGAAGGGGTA
TTTTGCCACTGGAGGAAAGTGAAGTCGACCTCCAGTACTCAAGAGACCAAGTAGCAGCTCTTGGCCACAGAGCC
CAGCCCCCAGACAACCTGAACTCAAATGAGATGCCGAGGACAGCAGCACTCCCGAAGAGATGCCACCTCCAG
AGCCCCCAGAGCCACCAAGGAGGACGCTGAAGCTGAGAAGTAGCCTATCTATGGAAGAGACTTTTGTTTGTGTT
TAATTAGGCGTATGAGAGATTTTCAAGGTGAGAAGTTAAACCTGAGACAGAGAGCAAGTAAGCTGTCCCTTTTAACT
GTTTTTCTTTGGTCTTTAGTCAACCCAGTTGCACTGGCATTTTCTGTGCAAGCTTTTAAATTTCTGAACT
CAAGGCAGTGGCAGAAGATGTGAGTCACTCTGTATACTGGAAAAATGGGTCTCTTGGGCCCTGGCACTGGTCTCT
CCATGGCCCTCAGCCACAGGGTCCCTTGGACCCCTCTCTTCCCTCCAGATCCAGCCCTCTCTGTCTGGGTGTCAC
TGTCTCTATTCTGGGCTAAAAGTTTTTGTAGACTGGCTCAAATCTCCTCCAGCTGTCTGCAAGTGTGAGTCCAGA
GGCAGTCAAGAGACCTCTGGCCAGGGGATCCTAACTGGGTCTTGGGGTCTTCAGGACTGAAGAGGAGGGAGAG
TGGGGTCAGAAGATCTCTTGGCCACCAAGTGCCAGCAITGCCCAAATCCTTTTAGGAATGGGACAGGTACCT
TCCACTTGTGTANNNNNNNNNNNNNNNNNNNNNNNNNNNNNTGTTTTTCTTTTGAATCTCTGCTCCCATTAGGAG
CAGGAATGGCAGTAAATAAAGTCTGCACCTTGGTCAITTTCTTTCTCAGAGGAAGCCCGAGTGCTCACTTAAAC
ACTATCCCTCAGACTCCTCTGTGTGAGGCTGCAGAGGCCCTGAATGCAAAATGGGAAACCAAGGCACAGAGAG
GCTCTCTCTCTCTCTCTCTCCCCGATGTACCTTCAAAAAAATAAGTCTAAACAGTCTTCTCTCAATTAAGCCT
CGGCTGAGTGAGGGAAGCCAGCACTGTGCTCCCTCTCGGGTAACTCACCTTAAGGCTCTCGGCCACTCTGCTCT
ATGGTAACCACTAGGGGCTTCTCTCAAGCCCGCTCTTCCAGCACTTCCACCGGCAGTCCCGGCCACTCTGCT
CACCTTGGGGTGGGCTGTGGCCCCCAGTCACTCTGCTCAGGACTGCTCTATTGAGGGAAGAAGATTATGT
ATTATATGTGGCTATATTTCTAGAGCAGCTGTGTTTCTCTTTCTTAAGCAGGGTCTCTGTGATGACTTAT
CGGCTGGGGAGGTAAACCGAACTTTTCACTATTGTAAGCGATTAAACTGTGTCTAATGCA

FIGURE 219

MSVMVVRKKVTRKWEKLPGRNTFCCDGRVMMARQKGIFYLTFLILGTCTLFFAFECRYLAV
QLSPAIPVFAAMLFLFSMATLLRTSFSDPGVIPRALPDEAAFIEMIEATNGAVPQGQRPPP
RIKNFQINNQIVKLKYCYTCKIFRPPRASHCSICDNCVERFDHHCWPVGNVCVKRNYRYFYFYL
FILSLSLLLTIYVFAFNIVYVALKSLKIGFLETCLKETPGTVLEVLCFFTLWSVVGLTGFTTF
LVALNQTTNEDIKGSWTGKNRVQNPYSHGNIVKNCCVLCGPLPPSVLDRRGILPLEESGSR
PPSTQETSSSLLPQSPAPTEHLNSNEMPEDSSTPEEMPPPEPPEPPEQEAEEAEK

Putative transmembrane domains:

amino acids 36-55 (type II TM), 65-84, 188-208, 229-245

100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212
2213
2214
2215
2216
2217
2218
2219
2220
2221
2222
2223
2224
2225
2226
2227
2228
2229
2230
2231
2232
2233
2234
2235
2236
2237
2238
2239
2240
2241
22

FIGURE 220

AAAACCCCTGTATTTTTTACAATGCAAATAGACAATNANCCTGGAGGTCTTTGAATTAGGTAT
TATAGGGATGGTGGGGTTGATTTTTNTTCCTGGAGGCTTTTGGCTTTGGACTCTCNCCTTTCT
CCCACAGAGCNCCTTCGACCATCACTGCCCCCTGGGTGGGGAATTGTGTTGGAAAGAGGAATA
CCGCTANTTCTACCTCTTCATCCTTTNTCTCTCCNCCTCACAATCTATGTCTTCGCCTTCA
ACATCGT

100
90
80
70
60
50
40
30
20
10
0
100
90
80
70
60
50
40
30
20
10
0

FIGURE 221

GTGTGTCCTTCAGCAAAACAGTGGATTAAATCTCCTTGACACAAGCTTGAGAGCAACACAA
TCTATCAGGAAAGAAAGAAAGAAAAAACCGAACCTGACAAAAAGAAGAAAAAGAAGA
AAAAAATCATGAAAAACATCCAGCCAAAAATGCACAATTCTATCTCTTGGGCAATCTTCAC
GGGGCTGGCTGCTCTGTGTCTCTTCCAAGGAGTGCCTGTCGAGCGGAGATGCCACCTTCC
CCAAAGCTATGGACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCTCAGTGTGCACTATT
GACAACCGGGTCACCCGGGTGGCTGGCTAAACCGCAGCACCATCCTCTATGCTGGGAATGA
CAAGTGGTGCCTGGATCCTCGCGTGGTCTTCTGAGCAACACCCAAACGCAGTACAGCATCG
AGATCCAGAACGTGGATGTGTATGACGAGGGCCCTTACACCTGCTCGGTGCAGACAGACAA
CACCCAAAGACCTCTAGGGTCCACCTCATTTGTGCAAGTATCTCCAAAATTTGTAGAGATTT
TTCAGATATCTCCATTAAATGAAGGAACAATATTAGCCTCACCTGCATAGCAACTGGTAGAC
CAGAGCCTACGGTTACTTGGAGACACATCTCTCCAAAGCGGTTGGCTTTGTGAGTGAAGAC
GAATACTTGAAATTCAGGGCATCACCCGGGAGCAGTCAGGGGACTACGAGTGCAGTGCCTC
CAATGACGTGGCCGCGCCCGTGGTACGGAGAGTAAAGGTACCGTGAACTATCCACCATACA
TTTCAGAAGCCAAGGGTACAGGTGTCCCCTGGGACAAAAGGGGACACTGCAGTGTGAAGCC
TCAGCAGTCCCTCAGCAGAATTCAGTGGTACAAGGATGACAAAAGACTGATTGAAGGAAA
GAAAGGGGTGAAAGTGGAAAAAGACCTTTCTCTCAAACTCATCTTCTCAATGTCTCTG
AACATGACTATGGGAATACTTGCCTGGCCCTCAACAAGCTGGGCCACACCAATGCCAGC
ATCATGCTATTTGGTCCAGCGCCCGTCAGCGAGGTGAGCAACGCGACGTGAGGAGGGCAGG
CTGCGTCTGGCTGCTGCCTCTTCTGGTCTTGACCTGCTTCTCAAAATTTGATGTGAGTGCC
ACTTCCCCACCCGGGAAAGGTGCGGCCACCACCACCACCAACAACAGCAATGGCAACAC
CGACAGCAACCAATCAGATATATACAAATGAAATTAGAAGAAACACAGCCTCATGGGACAGA
AATTTGAGGGAGGGGAACAAAGAATACTTTGGGGGAAAAGAGTTTTAAAAAGAAATTGAA
AATTGCCTTGACAGATATTTAGGTACAATGGAGTTTTCTTTTCCCAAACGGGAAGAACACAGC
ACACCCGGCTTGGACCACTGCAAGCTGCATCGTGCAACCTCTTTGGTGCCAGTGTGGGCAA
GGGCTCAGCCTCTCTGCCACAGAGTGCCCCACGTGGAACATTCTGGAGCTGGCCATCCCA
AATTCAATCAGTCCATAGAGACGAACAGAATGAGACCTTCCGGCCCAAGCGTGGCGCTGCGG
GCACCTTTGGTAGACTGTGCCACCACGGCGTGTGTTGTGAAACGTGAAATAAAAGAGCAAAA
AAAAA

FIGURE 222

MKTIQPKMHNSISWAIFTGLAALCLFQGVPVRSGDATFPKAMDNVTVRQGESATLRCTIDNR
VTRVAWLNIRSTILYAGNDKWCLDPRVLLSNTQTQYSIEIQNVVDVYDEGPYTCSVQTDNHPK
TSRVHLIVQVSPKIVEISSDISINEGNNISLTCIATGRPEPTVTWRHISPKAVGVFVSEDEYL
EIQGITREQSGDYECASNDVAAPVVRVKVTVNYPPYISEAKGTGVPVGQKGTLCQCEASAV
PSAEFQWYKDDKRLIEGKKGVKVENRPFLSKLIFFNVSEHDYGNVTCVASNKLGHTNASIML
FGPGAVSEVSNGTSSRRAGCVWLLPLLVLHLLKLF

Signal peptide:

amino acids 1-28

20
40
60
80
100
120
140
160
180
200
220
240
260
280
300
320
340
360
380
400
420
440
460
480
500
520
540
560
580
600
620
640
660
680
700
720
740
760
780
800
820
840
860
880
900
920
940
960
980
1000

FIGURE 223

GAAAAAAAAATCATGAAAACCATCCAGCCAAAAATGCACAATTCTATCTCTTGGGCAATCTTC
ACGGGGCTGGCTGCTCTGTGTCTCTTCCAAGGAGTGCCCGTGCGCAGCGGAGATGCCACCTT
CCCCAAAGCTATGGACAACGTGACGGTCCGGCAGGGGAGAGCGCCACCCTCAGGTGCACTA
TTGACAACCGGGTCACCCGGGTGGCCTGGCTAAACCGCAGCACCATCCTCTATGCTGGGAAT
GACAAGTGGTGCCTGGATCCTCGCGTGGTCCTTCTGAGCAACACCCAAACGCAGTACAGCAT
CGAGATCCAGAACGTGGATGTGTATGACGAGGGCCCTTACACCTGCTCGGTGCAGACAGACA
ACCACCCAAAGACCTCTAGGGTCCACCTCATTGTGCAAGTATCTCCCAAATTGTAGAGATT
TCTTCAGATATCTCCATTAATGAAGGGAACAATATTAGCCTCACCTGCATAGCAACTGGTAG
ACCAGAG

100
90
80
70
60
50
40
30
20
10
0
A
C
G
T
100
90
80
70
60
50
40
30
20
10
0
A
C
G
T

FIGURE 224

ATGGCTGGTGACGGCGGGGGCCGGGCAGGGGACCGGGGCCGGCCCGGGAGCGGGCCAGCTGCCGGGAGCCCTGA
 ATCACCGCTCGGCCGACTCCACATAGAAACGTCGCGCTGCAGGAGCTGGGAGCTGGCAGCAACGCTGGGATTCCAG
 AAGGGGCAAGACAGCTGTTAGGCTCACGCACGCAGCTGGAGCTGGTCTTAGCAGGTGCCTCTCTACTGCTGGCT
 GCATGCTTCTGGGCTGCCTTGGGCCCTAGGGGTCCAGTACCACAGAGACCATTCCACAGCACCCTGCCTTACA
 GAGGCTGCATTGAGTGGCTGAAAAATCCTGGAGTCCCTGGACCGAGGGGTGAGCCCTGTGAGGACTTTTAC
 CAGTTCTCTGTGGGGCTGGATTTCGGAGGAACCCCTGCCGATGGGCGTCTCGCTGGAACACCTTCAACAGC
 CTCTGGGACCAAAACAGGCCATCTGAAGCACCTGCTTGAAAAACACACCTTCAACTCCAGCAGTGAAGCTGAG
 CAGAAGACACAGCGCTTCTACCTATCTTGCCCTACAGGTGGAGCGCATTGAGGAGCTGGGAGCCAGCCACTGAGA
 GACCTCATTTGAGAAGATTGGTGGTTGGAACATTACGGGGCCCTGGGACCGAGCAACTTTATGGAGGTGTGAAG
 GCAGTAGCAGGACCTACAGGGCCACCCATTCTCAGCGCTACATCAGTGCCGACTCTAAGAGTTTCAACAGC
 AATGTTATCCAGGTGGACAGCTCTGGGCTCTTTCTGCCCTCTGGGATTACTACTTAAACAGAACTGCCAATGAG
 AAAGTGCTCACTGCTTATCTGATTATACATGGAGGAATCGGGGATGCTGCTGGGTGGGCGGCCACCTCCACGAGG
 GAGCAGATGCAGCAGGTGCTGGAGTTGGAGATACAGCTGGCCAAACATCACAGTGCCCCAGGACCCAGCGCGCGAC
 GAGGAGAAGATTACACAAAGATGAGCAITTCGGAGCTGCAGGCTCTGGCGCCTCCATGCACTGGCTTGAATTC
 CTGCTCTTTCTGCTGCACCAATTGGAGTTGAGTGAATCTGAGCCTGTGGTGGTATGGGATGATTAATTTGAG
 CAGGTGTGAGAGCTCATCAACCGCAGGAACCAAGCATCTGAACAATTACCTGATCTGGAACCTGGTGC AAAAG
 ACAACCTCAAGCCTGGAACGACGCTTTGAGTCTGCACAAGAGAAGCTGCTGGAGACCCCTCTATGGCACTAAGAAG
 TCTGTGTGTCGAGGTGGCAGACCTGCATCTCAAACCGGATGACGCCCTTGGCTTTGGCTTTGGGGTCACTCTTC
 GTGAAGGCCACGTTTGACCGGCCAAAGCAAAATTCAGAGGGGATGATCAGCGAAATCCGGACCGCATTTGAG
 GAGGCCCTGGGACAGCTGTTTGGATGGATGAGAAGACCCGCGAGGAGCAAGGAGAAAGCAGATGCCATCTAT
 GATATGATTGGTTTCCAGACTTTATCCTGGAGCCCAAGAGCTGGATGATGTTTATGACCGGATGAGAAATTTCT
 GAAGATTCTTTCTTCAAAACATGTTGAATTTGTACAACTTCTCTGCCAAGGTTATGGCTGACCACTCCGCAAG
 CCTCCACGCGAGACCAAGTGGAGCATGACCCCGCAGACAGTGAATGCCCTACTACCTTCCAAC TAAGAAATGAGATC
 GTCTTCCCGCTGGCATCTCTGAGGCCCTTCTATGCCCGCAACCCCAAGGCCCTGAACTCTCGGTGGCATC
 GGTGTGGTCATGGCCATGAGTTGACGATGCTTTGATGACCAAGGGCGCGATGATGACAAAGAGGGAACCTG
 CGGCCCTGGTGGCAGAATGAGTCCCTGGCAGCCTTTCGGAACCAACAGGCCCTGCATGGAGGAACAGTACAATCAA
 TACCAGGTCAATGGGAGAGGCTCAACGCGCGCAGACGCTGGGGAGAGAACTACTGACAAACGGGGGCTGAAG
 GCTGCCTACAATGCTTACAAGCATGGCTGAGAAAGCATGGGGAGGAGCAGCAACTGCCAGCCGTGGGGCTCACC
 AACCACAGCTCTTCTTCTGTGGGATTTCGCCAGGTGTGGTGTCTCGGTCCGACACACAGAGAGCTCTCACAGAGGG
 CTGGTGACCGACCCCAACAGCCCTGCCCGCTTCGCGTGTGGGCACTCTCTCAATCCTCCGTGACTTCTGTGG
 CACTTCGCTGCCCTGTGGCTCCCCATGAACCCAGGGCAGCTGTGTGAGGTGTGGTAGACCTGGATCAGGGGA
 GAAATGGCCAGCTGTACACAGACCTGGGGCAGCTCTCTGACAAAGCTGTTTGTCTCTTGGGTTGGGAGGAAGCAA
 ATGCAAGCTGGGCTGGGTCTAGTCCCTCCCGCCACAGGTGACATGAGTACAGACCTCTCTCAATACACCAATTG
 TGCTCTGCTTGGGGGTGCCCTGCTCCAGCAGAGCCCAACATTTCACTGTGACATCTTTCGTGTCACTCT
 GCTTGAAGAGGTCTGGGTGGGAGGCCAGTTCCCATAGGAAGAGTCTGCC

100
 200
 300
 400
 500
 600
 700
 800
 900
 1000
 1100
 1200
 1300
 1400
 1500
 1600
 1700
 1800
 1900
 2000
 2100
 2200
 2300
 2400
 2500
 2600
 2700
 2800
 2900
 3000
 3100
 3200
 3300
 3400
 3500
 3600
 3700
 3800
 3900
 4000
 4100
 4200
 4300
 4400
 4500
 4600
 4700
 4800
 4900
 5000
 5100
 5200
 5300
 5400
 5500
 5600
 5700
 5800
 5900
 6000
 6100
 6200
 6300
 6400
 6500
 6600
 6700
 6800
 6900
 7000
 7100
 7200
 7300
 7400
 7500
 7600
 7700
 7800
 7900
 8000
 8100
 8200
 8300
 8400
 8500
 8600
 8700
 8800
 8900
 9000
 9100
 9200
 9300
 9400
 9500
 9600
 9700
 9800
 9900
 10000

FIGURE 225

MNVALQELGAGSNVGFQKGTQQLLGSRTQLELVLAGASLLLAALLLGCLVALGVQYHRDP
STCLTEACIRVAGKILES LDRGVSPCEDFYQFSCGGWIRRNPLPDGRSRWNTFNSLWDQ
NQA ILKHLENTTFNSSSAEQKTQRFYLSCLQVERIEELGAQPLRDLIEKIGGWNITGPWDQ
DN FMEVLKAVAGTYRATPFPTVYISADSKSSNSNVIQVDQSGFLPSRDYILNRTANEKVL
TAY LDYMEELGMLLGG RPTSTREMQQVLELEIQ LANITVPQDQRRDEEKIYHKMSI
SELQALAP SMDWLEFLSFLLSPLELSDSEPVVVGMDYLQQVSELINRTEPSILNNYLI
WNLVQKTTSSL DRRFESAQEKLL ETLYGTTKSCVPRWQTCISNTDDALGFALGSLFV
KATFDRQSKEIAEGMI SEIRTA FEALGQLVWMD EKTQAAKEKADAIYDMIGFPDFI
LEPKELDDVYDGYEISED SF FQNMLNLYNFSAKVMADQLRKPPSRDQWSMT
PQTVNAYYLP TKNEIVFPAGILQAPFYARNH PKALNFGGIGVVMGHELTHAFDDQ
GREYDKEGNLRPWWQNESLA AFRNHTACMEEQYNQYQV NGERLNGRQTLGENITD
NGGLKAAYNAYKAWLRKHGEEQQLPAVGLTNHQLFFVGFAQVWCS VRTPESSHEGL
VTDPHSPARFRVLGTLSNSRDFLRHFGCPVGSPPMNPQLCEVW

Type II Transmembrane domain:

amino acids 32-57

[illegible]

CCGCGGCGCTTCGCGCCTTCGACTCCCGGCTCCCTCCCTCCGCGCGCTCCGCGCGCTCTCCTCCTCCCTCCCTCC
CAGCTGTCTCCGCTTCGCGTGCATCGCAGCGCTCCGCGCGCGCGCGCGCGCTGCTGTCTCCTCCGCGTGTCTGCTGCT
CGGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCTGCGCATCCGTTCTGAGAAAGAGCGCGTGTCTC
CGTTTCGGGAGCGCGCAGGTGAGTGGGCGCGTGGGCGGAGGCGCGGCGGGGAGTCGCGCTCGGCGCGAGTCAGCGCG
CAGCGCCGAGGGGCGCGCGGCGCGCAGGTGGCTCGGCGCGCGCGCGCGCGCGCGCGAGGGTGGGCGGGGCGAGAGGGC
CGCGTGTCTGGGAGCCGCGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGTGGGCGCGGCGCTCCAGC
CAAGCCCGCTCCCGCAGGCTGCACCTTCGCGGGAAGGTCTATGCTTGGAGCAGCTGGGCACTGGACCGCACTGAGG
GGAGCGCACTTCGGGCTGTGCGCTGTGCTGTGCGCTCGCAGCGCGCATGGGGTCCGCGCTCAGCAGGCGCGCTCGG
CAGGGTCACTCGAGAAACATCAACACAGAGTGGCCAAACCGCGCTGTGGGCGAGCGCGCAGCTGCGCGGACAC
CTGCTCGCAGACCTGCCCCAGCAGCTCTGTGCGCGTCTGACAGCGCGCGAGGTGCGAGCGCGGTGGCGAGCGCGC
AGTCTGCTGTGCGCTGTAGCTCCGCTCTCTATCTCTATGAGCGCGTGAGCCGCCCTACAGGATCCGCT
CTCAGACTCCAAATGGCAGTGTCTGTCTGTAGACAGCTCGACGCCACCCAGAGATGGCTGTGTGTGGGGTGTG
CGCGGCACTGCTCGGTTGTCTTCGGCGCTCTTAGGCGAAGACAGCTGCATGTGGCACTGTGACATCTCATCA
CCCTTCAGGCGAGGTCTGGGGGCTCTCATCCGCGCACCGGCGCTGTCCCGAGAGACCTTCAGTGCATCTGCATC
TCTAGAAGGCCCCACACGACGCGGCTGAGGAGGATCACTGCTCACTTCAGTGAACAGAGGATCCTTGTGCA
TTTTTTGTGCTCTTCGAGGCGTTCGAGGAGTAAACCGAGTTCCTTGAAGGCTCAGAGTTCTACACAGGGGCA
GCTACTGCGAGAACTTCGCGCAATGTCTCAGGCCAGGAAACCGAGGTTTGTGTGAGTGTGCGCCAACTGCTCAGT
CCAGAGAGAGCTAGCTGGCTGTGCTGCGGGAGCTGCAGATGGCCCTGGAAGTGGGACGAGCGCGCGCTGCGCAT
CAGTGAAGCAATCTGCTGCCAAGAGAGAGCTGCACGCTCTGCTGCAAAAGTGTCTTTTGTGGGGCTAATGCTCCGATCCC
AGTCCAAACCGGTCGTCGCGCTCGACGAGCCTCACTCTGTGAGAAATGGGNCCTTGTTCCTCAGGTGCAATTT
GGTAGGGAACAACAGTGAAGTGGTGGGCGCATGACATGAAACCAAGCTCAGCGGAGGATCAGCCCACTGCTCTCT
GTGCAAGGCTGGGCTGGCTATCTCCCTGCCCGAGCGCGTGGGTATCTGCGCTGGCTGGGGTGGCCGAGGGGG
TCATATGTGCTGTCGAGAAATGAGCTCTCTCTGAGCTGGGACACAGGAGACTCCGACGAGGAGCTCTCGGGGGCA
AGTGTGCTGCTCTGCCCTACTGTGGGGATGAGCGCGCGCTGCGTCCGCTAGCAGAGAGCTCTGTGTGTACTCC
CCTGTGTGAAGAGCCAAAGCAGCGAGGACGAGCTGGCTTCTCTTGGATACCACTCTGTACCTGTCACTATGAAGTGT
CTGGCTGGCTTGGTGGTCTCAGAAACAGGCACTGTACCTGCCACCTCTTGGCGCTCTGGAACGCCAGGGCC
TCGCGGGCTGTGAAGAGGATCTATGTGCTCAGAGGCCAGGGTGTGTGAAAGAACCTTGAGAGCCGAGACTGCTGCG
GCACCTCGGAAAGAGGCTAGCTTCCCTGATGATACCAACCAAGTGAACCCGAGGGGAGCTCCAGGAGAGCC
CTCTCCACAGGTGCACATAGCCAAACCAATGTGAGGTTGGCGGACTGCGCTGGAGGCGGCGGGGCGAGGGGGT
CGGGCGCTCGGGGCGCTCCGATACAGCTCTGTGTCGCGCGCTGTGTGTGCTGTGTCTCCCGGCTCTAGCGCCCGAC
CAAACCTGTGTGTCTGGGCGCGCCCGAGAACCCAAACAGTGTCTTCGAGGGGAGCAGCGCCCAACGAGGCG
TGCTCGGGCGCCCAACTGACGACCGGCTCTGCTCACTGTGCACCTGCGACAGTGCAGAGAGCAGCGGTGATCTGTGACCGGCT
GGTGTGGCCACCGCGCAGCTGCCCAACCGCGCTGCGAGCTCCGACAGTGTGCTGCTTGTTCCTGGCTGTGCTA
TTTGTGTGAGTGCAGCGGAGCTCGGGCGCGCGGGTACCGGGTACCGGTCGCGACCCGCTGTGCCCCCTTTGGCTTAATTA
TGTGTGCTCTGTGCATCTGCAGCAGGGGGGACATGAGAGAGTGACATCTGAGAAGGTGCAGTGTCCCGCGTGC
CTGTGCCAGCGCTGTGCTGTCTCAACCCACCGACTCTG2AAACAGATGTCCAGGTGAGGCGCACCCCGACGCTGGG
GAGCCCACTGACAGCTGTATGGGCGCGGGGCTCGGTTTGTCTGGGAGTGTGTCCACAGAGTCAAGAGCTGCGCA
CCCCCTAGTGCCCCCCTTTGAGAGATGATGAGTGTACCTCGCAGATGTGGGGTGGTGGGAGCGAGAGGCTGTGT
GTAGAGTGGGTAATGGGAGCGCTGTGTGTGGAGTAGGGAAGCTCCAGAGGAGTCTCTGAAGAAGCTCAAGGCT
GTGTGTCCAGTGCCTCGGGGACACATCAGTGTCTGTCTGTCTGTGTTGAGTGCAGAGGGGTGCCCTCACTGTGAGC
CGGATGACTGTCTACCTGCACTGTCTCTGTGCTGGCGGAAGGAGTGCATGCTGTCTCCGCTCAGCGGCCAC
GGCGCGTCAAGTGAAGGAGGTCCAGGCTGACAGCGTGTGAGTGTGAGGGTCACTGCCCTGTGGGACTCTGTATGAC
GAGAGGGAGCACTCATGTGTGCGAGAACAGTGCAGCCTGTGATCCAACTGCCATTCCAATCCACCTCAGCAGCA
ACTGTGTGGAATGTATTGTATTTGCTTTTCTTACAATGAGATTTCTGAAGTGCAGAGAAATTAAGCAACGAG
ATAAGAGTCAACCACTGTGTGCATGACCTGTGTGTAAGAAATCTGGCCTTCTGGGACAGAGGAGGAGTGTCT
TGCCCTGCCCTCTATGCGCTCTGTGTGCTCTGCATCCCTCTCCCTCTCCCAACTCTCCCTCTCTGTCTCT
AGACGCCAGAGAGCAGCAACTGTACTCAGAGCTGGAGAAAGGCGAGGCGCTTAGGAGGAGCGACGAGAGGGCC
AAGTGACCAAGAGAGTGGGCTCGAGCTGGGAAGGGGTGTGCATCGAGAGACTCTTGTCATTCTCTGTGGAAAG
CCGAGTGTGCTTGTGCTCTGTGCTGCTCTACTCCACCCCACTACTCTGGGAACACAGCTCTCAAGAGAGGGS
GAGAGGAGCTGGGCGAGACGAGGTACAGCGATCTCAGTGTCTGCCCTGCCACCTCGGCTCTCTCTGTGAAG
GCCCAACCGCTTTCTCTCTGTACATATGTCACTGGCTTTGTGGGATTTTAAATCTTATCTCACTCAGCAGCAAG
GGCCCGGACCACTCACTCTGCTGCCCTGAGCTGAGCAGAGCATTAATGTGGAGAGTTTGTATTTATAAACC
ATTTCTTTTTCAGTCTTTGGGCACTGAGTGTGCTCTTTGTGGCGAGGAACCTGAGTGGGGCGTGGTGGAGAGAGG
CGNAGAGATGAGAGGTGAGAGAGAGAGCTGTGACATCTGGGAGCTGAAGAGACTCGAGAGAGCAGAGATAGG
CTGTGCGNNTGGCTGGCATCCTCGGTTCCGAGAGGGGCTGGGGAGTGTGTCTTGAGATGTGTCTAGAGACTCAAG
AATTTAGAGTGTGCTCAGAGAGAGATTTTGAATCAAGTTTAGTTTCCACATCGCTGGCGCTTTTGTGACTTCACT
TTTGAAGTGTCTCCAGGAGAGATCAAGGTTGCACGAGCCCTCTCTCCCTCCTTCCCTTCCCTTCCCTTTCT
TTCCCTCCCTCCCTCCGCGCGCTCCCTCC

FIGURE 227

GGCCGAGCGGGGTGCTGCGCGGCGGCGGTGATGGCTGGTGACGGCGGGGCCGGGCAGGGGA
CCGGGGCCGCGGCGCGGGAGCGGGCCAGCTGCCGGGAGCCCTGAATCACCGCTGGCCCGAC
TCCACCATGAACGTGCGCTGCAGGAGCTGGGAGCTGGCAGCAACGTGGGATTCCAGAAGGG
GACAAAGACAGCTGTTAGGCTCACGCACGCAGCTGGAGCTGGTCTTAGCAGGTGCCTCTCTAC
TGCTGGCTGCACTGCTTCTGGGCTGCCTTGTGGCCCTAGGGGTCCAGTACCACAGAGACCCA
TCCCACAGCACCTGCCTTACAGAGGCTGCATTTCGAGTGGCTGGAAAAATCCTGGAGTCCCT
GGACCGAGGGGTGAGCCCTGTGAGGACTTTTACCAGTTCTCCTGTGGGGGCTGGATTTCGGA
GGAACCCCTGCCCCGATGGGCGTTCTCGCTGGAACACCTTCAACAGCCTCTGGGACCAAAC
CAGGCCATACTGAAGCACCTGCTTGAAAAACACCCTTCAACTCCAGCAGTGAAGCTGAGCA
GAAGACACAGCGCTTCTACCTATCTTGCTACAGGTGGAGCGCATTGAGGAGCTGGGAGCCC
AGCCACTGAGAGACCTCATTTGAGAAGATTGGTGGTTGGAACATTACGGGGCCCTGGGACCA
GACAACTTTATGGAGGTGTTGAAGGCAGTAGCAGGGACCTACAGGGCCACCCATTCTTCAC
CGTCTACATCAGTGCCGACTCTAAGAGTTCACACAGCAATGTTATCCAGGTGGACCACTG
GGCTCTTTCTGCCCTCTCGGGATTACTACTTAACAGAACTGCCAATGAGAAAAGTAAGGAAC
ATCTTCCGAACCCCATCCCTACCCCTGGCTGAGCTGGGCTGATCCCTGTTGACTTTTCCCT
TTGCCAAGGGTCAGAGCAGGGAAGGTGAGCCTATCCTGTACCTAGTGAACAAACTGCCCT
CCTTTCTTTCTTTCTTTCTTCTCCTCCCTCCCTCCCTTTCTTCCCCTTTTCTTCTTCTTCTC
TCTTATTCTTCTAGTAGGTTTCATAGACACCTACTGTGTGCCAGGTCCAGTGGGGGAATTG
GAGATATAAGTTTCCGAGCCATTGCCACAGGAAGCGTTCAGTGTGCATGGGTTTCATGGACCT
AGATAGGCTGATAACAAAGCTCACAGAGGGTCTGAGGATTGAGGAGAGACTTATGGAGCC
AGCAAAGTCTTCTCTGAAGAGATTGCATTTGAGCCAGGTCTGTAG

FIGURE 228

ATGCCTACTACCTTCCAATAAGAATGAGATCGTCTTCCCCGCTGGCATCCTGCAGGCCCCC
TTCTATGCCCGCAACCAACCCCAAGGCCCTGAACCTTCGGTGGCATCGGTGTGGTCATGGGCCA
TGAGTTGACGCATGCCTTTGATGACCAAGGGCGCGAGTATGACAAAGAAGGGAACCTGCGGC
CCTGTGGCAGAAATGAGTCCCTGGCAGCCTTCCGGAACACACGGCCTGCATGGAGGAACAG
TACAATCAATACCAGGTCAATGGGGAGAGGCTCAACGGCCGCCAGACGCTGGGGGAGAACAT
TGCTGACAAACGGGGGGCTGAAGGCTGCCTACAATGCTTACAAGCATGGCTGAGAAAGCATG
GGGAGGAGCAGCAACTGCCAGCCGTGGGGCTCACCAACCACCAGCTCTTCTTCGTGGGATTT
GCCCAGGTGTGGTGCTCGGTCCGCACACCAGAGAGCTCTCACGAGGGGCTGGTGACCGACCC
CCACAGCCCTGCCCGCTTCCGCGTGTGGGCACTCTCTCCAACCTCCCGTGACTTCCTGCGGC
ACTTCGGCTGCCCTGTGCGGTCCCCCATGAACCCAGGGCAGCTGTGTGAGGTGTGGTAGACC
TGGATCAGGGGAGAAATGGCCAGCTGTCAACAGACCTGGGGCAGCTCTCTGACAAAGCTGT
TTGCTCTTGGGTGGGAGGAAGCAAATGCAAGCTGGGCTGGGTCTAGTCCCTCCCCCACA
GGTGACATGAGTACAGACCCTCCTCAATCACCACATTGTGCCTCTGCTTTGGGGGTGCCCT
GCCTCCAGCAGAGCCCCACCATTCAGTGTGACATCTTCCGTGTCACCTGCCTGGAAGAG
GTCTGGGTGGGGAGGCCAGTTCCCATAGGAAGGAGTCTGCCTCTTCTGTCCCCAGGCTCACT
CAGCCTGGCGCCATGGGGCTGCCGTGCCTGCCCACTGTGACCCACAGGCCCTGGGTGGTG
TACCTCCTGGACTTCTCCCCAGGCTCACTCAGTGCGCACTTAGGGGTGGACTCAGCTCTGTC
TGGCTCACCTCACGGGCTACCCCCACCTCACCTGTGCTCCTTGTGCCACTGCTCCCACTG
CTGCTGCTGACCTTCACTGACAGCTCCTAGTGGAAGCCCAAGGGCCTCTGAAAGCCTCCTGC
TGCCCACTGTTTCCCTGGGCTGAGAGGGGAAGTGCATATGTGTAGCGGTACTGGTTCTGT
GTCTTAGGGCACAAGCCTTAGCAAATGATTGATTCTCCCTGGACAAAGCAGGAAAGCAGATA
GAGCAGGGAAAAGGAAGAACAGAGTTTATTTTACAGAAAAGAGGTTGGGAGGTTGTGGTCT
TGGCCCTTATAGGACC

FIGURE 229

CCCACGCGTCCGAGCCGCCGAGAAATTAGACACACTCCGACGCGGCCAAAAGCAACCGAGA
GGAGGGGAGGCAAAAACCCGAAAAACAAAAGAGAGAAACCAACCCCAACTGGGGTGG
GGGGAGAAAGAAAGAAAAGAAACCCACCACCCCAAAAAAATAAAAAA
AAAAAATAAAATCCTGTGGCGCGCCCTGGTTCCCGGAAGACTCGCCAGCACCAGGGGG
TGGGGGAGTGCGAGCTGAAAGCTGCTGGAGAGTGAGCAGCCCTAGCAGGGATGGACATGATG
CTGTTGGTGCAGGGTGCTTGTGTCTGAACCACTGGCTGGCGGGCGGTGCTCCTCAGCTGTG
CTGCCTGCTACCTCTCCTGCTCCCGCTGGACAGAGTGTGGAATTTCCCTGGGCGGCCGTGG
ACAACATGATGCTCAGAAAAGGGGACACGGCGGTGCTTAGGTGTTATTTGGAAGATGGAGCT
TCAAAGGGTGCCCTGGCTGAAACCGGTCAAGTATTATTTTGGCGGAGGTGATAAGTGGTCAGT
GGATCCTCGAGTTTCAATTTCAACATTGAATAAAAGGGACTACAGCCTCCAGATACAGAATG
TAGATGTGACAGATGATGGCCCATACACGTGTTCTGTTGAGACTCAACATACACCCAGAACA
ATGCAAGGTGCATCTAACTGTGCAAGTTCCTCCTAAGATATATGACATCTCAAATGATATGAC
CGTCAATGAAGGAACCAACGTCACCTTACTTGTGGCCACTGGGAAACAGAGCCTTCCA
TTTCTTGGCGACACATCTCCCATCAGCAAAACCAATTTGAAAATGGACAATATTTGGACATT
TATGGAATTACAAGGGACCAGGCTGGGAATATGAATGCAGTGCAGAAAATGCTGTGTCAAT
CCCAGATGTGAGGAAGTAAAAGTTGTTGTCACTTTGTCTCCTATTTGAGGAAATTAAT
CTGGCACCCTGACCCCGGACGCACTGGCCCTGATAAGATGTGAAGGTGCAGGTGTGCGCCCT
CCAGCCTTTGAATGGTACAAGGAGAGAAGAAGCTCTTCAATGGCCAAAGGAATATTATT
TCAAAATTTTAGCAACAGATCCATTCTCACTGTTACCAACGTGACACAGGAGCACTTCGGCA
ATTATACCTGTGTGGCTGCCAACAAAGCTAGGCACAAACCAATGCGAGCCCTGCTCTTAACCT
CTAGGTACAGCCAGTATGGAATTACCGGGAGCGCTGATGTTCTTTCTCCTGCTGTGTACCT
TCTGTTGACACTGTCTCTTTCCACAGCATATTTCTACCTGGAAGATGCCATTCTCAATAAA
TTCAAAGACCCATAAAAGGCTTTTAAGGATTCTCTGAAAGTGCTGATGGCTGGATCCAATCT
GGTACAGTTTGTAAAAGCAGCGTGGGATATAATCAGCAGTGCTTACATGGGGATGATCGCC
TTCTGTAGAATTGCTCATTATGTAATACTTTAATTCTACTCTTTTGGATTAGCTACATTA
CCTTGTGAAGCAGTACACATTTGCTCTTTTAAAGACGTGAAAGCTCTGAAATTACTTTTAG
AGGATATTAAATTTGTGATTTTCATGTTTGTAAATCTACAACCTTTCAAAGCATTCAGTCATGGT
CTGCTAGGTTGCGGCTGTAGTTTACAAAAACGAATATTGCAAGTGAATGTGATCTTTTAA
GGCTGCAATACAAGCATTCAGTTCCTCTGTTTCAATAAGAGTCAATCCACATTTACAAGATG
CATTTTTTTCTTTTGTATAAAAAAGCAAATAATATTGCTTCAGATTATTTCTTCAAATA
TAACACATATCTAGATTTTCTGCTTGATGATATTACAGTTCAGGAATGAGCCTTGTAAT
ATAACTGGCTGTGCAGCTCTGCTTCTCTTCCGTGAAGTTGAGCATGGGTGTGCCCTTCATAC
AATAATATTTTTCTCTTGTCTCCAACATAATAAAATGTTTGTCAAACTTACAATTTGA
AAGTAAAAATAAACAGAGATGATCAAGTTAAACCATACATATCTCTAAGTAACGAAGGAGC
TATTGGACTGTAAAAATCTCTTCTGCACTGCAATGGGGTTGAGAAATTTGCCCCACACT
AACTCAGTTCCTGTGATGAGAGACAATTTAATAACAGTATAGTAAATATACCATATGATTTT
TTAGTTGTAGCTAAATGTTAGATCCACGTTGGGAAATCATTCCCTTTAAATGACAGCACA
GTCCACTCAAAGGATTGCTGATCAAGATCATCTTTCTTCTACTAGTCCAAGCAAAAA
TTTTAAGATGATTTGTGAGAAAGGCACAAAGTCCATCACCTAATATTACAAGAGTTGGTA
AGCGCTCATCATTAAATTTTATTTTGTGGCAGGTATTATGACAGTCGACTGGAGGGTATGGA
TATGGATATGGACGTTCCAGAGACTATAATGCGAGAAACAGGTTAGGTGATCCGCTACTC
AGGAGGAAATTACAGAGACAAATTATGACAACGTGAAATGAGACATGCACATAATATAGATACA
CAAGGAATAATTTCTGATCCAGGATCGTCTTCCAAATGGCTGATTATATAAAGTTTTCG
AGCTGCACTGAAGCATCTTATTTTATAGTATATCAACCTTTGTTTTTAAATTTTACCTGCCA
AGGTAGCTGAAGACCTTTTAGACAGTTCATCTTTTTTTTTAAATTTTTTCTGCTTATTTAA
AGACAAATATTGGGAGCTTTGTCAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 230

MMLLVQGACCSNQWLA AVL LLSLCCLLPSCLPAGQSVDFPWA AVDNMMVRKGD TAVLR CYLED
GASKGAWLNRSSII FAGGDKWSVDPRVSISTLNKRDSLQIQNV DVTDDGPYTCSVQTQHTP
RTMQVHLTVQVPPKIYDISNDMTVNEGTVNLTCLATGKPEPSISWRHISPSAKPFENGQYL
DIYGITRDQAGEYEC SAENAVSFDPVRKVKVVVNFAPTIQEIKSGTVTPGRSGLIRCEGAGV
PPPAFEWYKGEKKLFNGQQGII IQNFSTRSILTVTNVTQEHFGNYTCVAANKLGT TNASLPL
NPPSTAQYGITGSADVL FSCWYLVLTLS SFTSIFYLKNAILQ

Important features of the protein:

Signal peptide:

amino acids 1-31

Transmembrane domain:

amino acids 326-345

N-glycosylation sites.

amino acids 71-75, 153-157, 273-277, 284-288, 292-296, 305-309

Casein kinase II phosphorylation site.

amino acids 147-151, 208-212, 224-228

Tyrosine kinase phosphorylation site.

amino acids 178-186

N-myristoylation sites.

amino acids 7-13, 63-70, 67-73, 151-157, 239-245, 291-297,
302-308, 319-325

Myelin P0 protein:

amino acids 92-121

FIGURE 231

AGTGGTTCGATGGGAAGGATCTTTCTCCAAGTGGTTCCTCTTGAGGGGAGCATTTCTGCTGG
CTCCAGGACTTTGGCCATCTATAAAGCTTGGCAATGAGAAATAAGAAAATTTCTCAAGGAGGA
CGAGCTCTTGAGTGAGACCCAAAGCTGCTTTTACCAAATTGCAATGGAGCCTTTTCGAAA
TCAATGTTCCAAAGCCCAAGAGGAGAAATGGGGTGAACCTTCTCCCTAGCTGTGGTGGTCATC
TACCTGATCCTGCTCACCGCTGGCGCTGGGCTGCTGGTGGTCCAAGTTCTGAATCTGCAGGC
GCGGCTCCGGTCTTGAGATGTATTTCTCAATGACACTCTGGCGGCTGAGGACAGCCCGT
CCTTCTCCTTGCTGCAGTCAGCACACCCTGGAGAACCTGGCTCAGGGTGCATCGAGGCTG
CAAGTCTCTGCAGGCCCAACTCACCTGGGTCCGCGTCAGCCATGAGCACTTGCTGCAGCGGGT
AGACAACCTCACTCAGAACCAGGGATGTTCAGAATCAAAGGTGAACAAGGCGCCCCAGGTC
TTCAAGGTCAACAAGGGGGCCATGGGCATGCCTGGTGGCCCCCTGGCCCCCGGGGACCCTGCT
GAGAAGGGAGCCAAGGGGGCTATGGGACGAGATGGAGCAACAGGCCCCCTCGGGACCCCAAGG
CCCACCGGAGTCAAGGGAGAGGCGGGCCTCCAAGGACCCAGGGTGCTCCAGGGAAGCAAG
GAGCCACTGGCACCCAGGACCCCAAGGAGAGAAGGGCAGCAAAGGCGATGGGGGTCTCATT
GGCCCAAAGGGGAACTGGAACTAAGGGAGAGAAAGGAGACCTGGGTCTCCAGGAAGCAA
AGGGACAGGGGCATGAAAGGAGATGCAGGGGTCTATGGGGCTCCTTGAGCCAGGGGAGTA
AAGGTGACTTCGGGAGGCCAGGCCACCAGGTTTGGCTGGTTTTCTTGAGCTAAAGGAGAT
CAAGGACAACCTGGACTGCAGGGTGTTCGGGGCCTCCTGGTGCAGTGGGACACCAGGTGC
CAAGGGTGAGCCTGGCAGTGTGGCTCCCCTGGGCGAGCAGGACTTCCAGGGAGCCCCGGA
GTCCAGGAGCCACAGGCCTGAAAGGAAGCAAAGGGGACACAGGACTTCAAGGACAGCAAGGA
AGAAAAGGAGAATCAGGAGTTCAGGCCCTGCAGGTGTGAAGGGAGAACAGGGAGCCAGG
GCTGGCAGGTCCAAGGGAGCCCCCTGGACAAGCTGGCCAGAAAGGAGACCAGGGAGTGAAAG
GATCTTCTGGGGAGCAAGGAGTAAGGGAGAAAAAGGTGAAAGAGGTGAAAACTCAGTGTC
GTCAGGATTGTGGCAGTAGTAACCGAGGCCGGGCTGAAGTTTACTACAGTGGTACCTGGGG
GACAATTTGCGATGACGAGTGGCAAATTTCTGATGCCATTGTCTTCTGCCGATGCTGGGTT
ACTCCAAAGGAAGGGCCCTGTACAAAGTGGGAGCTGGCACTGGGCAGATCTGGCTGGATAAT
GTTCAAGTGTGGGGCACGGAGAGTACCTGTGGAGCTGCACCAAGAATAGCTGGGGCCATCA
TGACTGCAGCCACGAGGAGGACGAGGCGTGGAGTGAGCGCTGGAACCGGAAACCTTTCA
CTTCTCTGCTCCCGAGGTGTCTCGGGCTCATATGTGGGAAGGCAGAGGATCTCTGAGGAGT
TCCCTGGGGACAACCTGAGCAGCCTCTGGAGAGGGGCCATTAATAAAGCTCAACATCATTGA

FIGURE 232

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA68886

><subunit 1 of 1, 520 aa, 1 stop

><MW: 52658, pI: 9.16, NX(S/T): 3

MRNKKILKDELLSETQQAAPHQIAMEPFEINVPKPKRRNGVNFSLAVVVIYLILLTAGAGL
LVVQVLNLQARLRVLEMYFLNDTLAAEDSPFSLLQSAHPGEHLAQGASRLQVLQAQLTWVR
VSHEHLLQRVDNFTONPGMFRIKGEQGAPGLQGHKGAMGMPGAPGPPGPPAEKGAKGAMGRD
GATGPGSPQGPFPVKGEAGLQGPQGPAGKQATGTPGPQGEKGSKGDGGLIGPKGETGTKGE
KGDGLPGSKGDRGMKGDAGVMGPPGAQGSKGDGFRPGPPGLAGFPGAQKGDQGGPGLQGVPG
PPGAVGHHPGAKGEPGSAGSPGRAGLPSPGSPGATGLKGSKGDGTLQGGQQGRKGESGVFGPA
GVKGEQGSPLAGPKGAPQAGQKGDQGVKGSSEGEQGVKGEKGERGENSVSVRIVGSSNRGR
AEVYYSGTWGTICDDEWQNSDAIVFCRMLGYSKGRALYKVAGAGTQQIWLDDNVQCRGTSTLW
SCTKNSWGHHDSCSHEEDAGVECSV

Transmembrane domain:

amino acids 47-66 (type II)

N-glycosylation sites.

amino acids 43-47, 83-87, 136-140

Tyrosine kinase phosphorylation site.

amino acids 432-440

N-myristoylation sites.

amino acids 41-47, 178-184, 253-259, 274-280, 340-346, 346-352,
400-406, 441-447, 475-481, 490-496, 515-521

Amidation site.

amino acids 360-364

Leucine zipper pattern.

amino acids 56-78

Speract receptor repeat

amino acids 422-471, 488-519

Clq domain proteins.

amino acids 151-184, 301-334, 316-349

FIGURE 233

CCCACGCGTCCGAAGGCAGACAAAGGTTTCATTTGTAAAGAAGCTCCTTCCAGCACCTCCTCT
CTTCTCCTTTTGCCCAAACCTACCCAGTGAGTGTGAGCATTTAAGAAGCATCCTCTGCCAAG
ACCAAAAGGAAAGAAGAAAAAGGGCCAAAAGCCAAAATGAAACTGATGGTACTTGTTTTTAC
CATTGGGCTAACTTTGCTGCTAGGAGTTCAAGCCATGCCTGCAAAATCGCCTCTCTTGCTACA
GAAAGATACTAAAAGATCACAACCTGTACAACCTTCCGGAAGGAGTAGCTGACCTGACACAG
ATTGATGTCAATGTCCAGGATCATTTCTGGGATGGGAAGGGATGTGAGATGATCTGTTACTG
CAACTTCAGCGAATTGCTCTGCTGCCAAAAGACGTTTTCTTTGGACCAAGATCTCTTTTCG
TGATTCTTGCAACAATCAATGAGAATCTTCATGTATTCTGGAGAACACCATTCTCTGATTTTC
CCACAACTGCACTACATCAGTATAACTGCATTTCTAGTTTCTATATAGTGCAATAGAGCAT
AGATTCTATAAATCTTACTTGTCTAAGACAAGTAAATCTGTGTTAAACAAGTAGTAATAAA
AGTTAATTCAATCTAAAAA

100
90
80
70
60
50
40
30
20
10
0
A
C
G
T
-
100
90
80
70
60
50
40
30
20
10
0
A
C
G
T
-
100
90
80
70
60
50
40
30
20
10
0
A
C
G
T
-
100
90
80
70
60
50
40
30
20
10
0
A
C
G
T
-

FIGURE 234

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52758
<subunit 1 of 1, 98 aa, 1 stop
<MW: 11081, pI: 6.68, NX(S/T): 1
MKLMVLVFTIGLTL LLGVQAMPANRLSCYRKILKDHNCHNLPEGVADLTQIDVNVQDHFWDG
KGCEMICYCNFSELLCCPKDVFFGPKISFVIPCNNQ

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 72-76

Tyrosine kinase phosphorylation site.

amino acids 63-71

FIGURE 235

CCCACGCGTCCGCGGACGCGTGGGCTGGACCCAGGCTCTGGAGCGAATTCAGCCTGCAGGG
 CTGATAAGCGAGGCGATTAGTGAGATTGAGAGAGACTTTACCCGCGCGTGGTGGTGGAGGGC
 GCGCAGTAGAGCAGCACACAGGCGCGGGTCCCGGAGGCGCGGCTCTGCTCGCGCGAGATG
 TGGAAATCTCCTTCACGAAACCGACTCGGCTGTGGCCACCGCGCGCCGCGCTGGCTGTG
 CGCTGGGCGCGTGGTGCTGGCGGGTGGCTTCTTTCTCCTCGGCTTCTCTTCGCGGTGGTTTA
 TAAAATCCTCCAATGAAGCTACTAACATTACTCCAAAGCATAATATGAAGCATTTTGGAT
 GAATTGAAAGCTGAGAACATCAAGAAGTCTTTACATAATTTTACACAGATACCACATTTAGC
 AGGAACAGAACAAAATTTAGCTTTGCAAAGCAAATTCATCCAGTGGAAAGAATTTGGCC
 TGGATTCTGTTGAGCTAGCTCATTATGATGTCTGTTGCTTACCCAAAATAAGACTCATCCC
 AACTACATCTCAATAATTAATGAAGATGGAAATGAGATTTTCAACACATCATTATTTGAACC
 ACCTCCTCCAGGATATGAAAATGTTTCGGATATTGTACCACCTTTTCAGTGCCTTTCTCTCCTC
 AAGGAATGCCAGAGGGCGATCTAGTGTATGTTAACTATGCCAGAACTGAAGACTTCTTTAAA
 TTGGAACGGGACATGAAAATCAATTGCTCTGGGAAAATTTGAATTTGCCAGATATGGGAAAAT
 TTTTCAGAGGAAATAAGGTTAAAAATGCCAGCTGGCAGGGGCCAAAGGAGTCACTCTCTACT
 CCGACCTGCTGACTACTTTTCTCCTGGGGTGAAGTCTTATCCAGACGGTTGGAATCTTCTCCT
 GGAGGTGGTGCTCCAGCGTGGAAAATATCCTAAATCTGAATGGTGAGGAGACCTCTCACACC
 AGGTTACCCAGCAAATGAATACTGCTTATAGGCGTGAATTTGCAAGGAGTGTGGTCTTCCAA
 GTATTCTGTTTCAATGGATACTATGATGCACAGAAGTCTCTAGAAAAAATGGGTGGC
 TCAGCACCACAGATAGCAGCTGGAGAGGAAGTCTCAAAGTGCCCTACAATGTTGGACCTGG
 CTTTACTGGAAAATTTTCTACACAAAAAGTCAAGATGCACATCCACTCTACCATGAAGTGA
 CGAGAATTTACAATGTTGATAGGTACTCTCAGAGGAGCAGTGGAAACAGACAGATATGTCAAT
 CTGGGAGGTCAACCGGACTCATGGGTGTTTGGTGGTATTGACCTCAGATGGGAGCAGCTGT
 TGTTTCATGAAATTTGTGAGGAGCTTTGGAACACTGAAAAGGAAGGGTGGAGACCTAGAAGAA
 CAATTTTGTGTTCAAGCTGGGATGCAGAGAATTTGGTCTTCTTGGTCTTACTGAGTGGGCA
 GAGGAGAATTCAGACTCCTTCAAGAGCGTGGCGTGGCTTATATTAATGCTGACTCATCTAT
 AGAAGGAAACTACACTCTGAGAGTTGATTTGTACACCGCTGATGTACAGCTTGGTACACAACC
 TAACAAAAGAGCTGAAAAGCCCTGATGAAGGCTTTGAAGGCAAATCTCTTTATGAAGTTGG
 ACTAAAAAAGTCTCTCCAGAGTTTCAAGTGGCATGCCAGGATAAGCAAATTTGGGATCTGG
 AAATGATTTTGAAGTGTCTTCCAAAGCTTGGAAATTTGCTTCAGGCAGAGCAGCGGTATACTA
 AAAATTTGGGAAACAAACAAATTCAGCGGCTATCCACTGTATCACAGTGTCTATGAACATAT
 GAGTTGGTGGAAAAGTTTTATGATCCAATGTTTAAATATCACCTCACTGTGGCCAGGTTTCG
 AGGAGGGATGGTGTGTTGAGCTAGCCAATTCATAGTGCTCCCTTTTGATTTGTGCGAGATTATG
 CTGTAGTTTTTAAGAAAAGTATGCTGACAAAATCTACAGTATTTCTATGAAACATCCACAGGAA
 ATGAAGACATACAGTGTATCAATTTGATTCACTTTTTTCTGCAAGTAAAGAAATTTACAGAAAT
 TGCTTCCAAGTTTCAGTGAGAGACTCCAGGACTTTGACAAAAGCAACCAATAGTATTAAGAA
 TGATGAATGATCAACTCATGTTTCTGAAAAGAGCATTATTATGATCCATTAGGGTTACCCAGAC
 AGGCCCTTTTATAGGCATGTCATCTATGCTCCAAGCAGCCACAACAAGTATGCAGGGGAGTC
 ATTTCCAGGAATTTATGATGCTCTGTTTGTATATTGAAAGCAAAGTGGACCTTCTCAAGGCCT
 GGGGAGAAGTGAAGAGACAGATTTATGTTGTCAGCCTTTCACAGTGCAGGCAGCTGCAGAGACT
 TTGAGTGAAGTAGCCCTAAGAGGATTTTTTAGAGAATCCGTATTGAATTTGTGTGGTATGTCA
 CTCAGAAAAGAAATCGTAATGGGTATATTGATAAATTTTAAAATTTGGTATATTTGAAATAAAGT
 TGAATATTATATATAA

FIGURE 236

```
></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA52756
><subunit 1 of 1, 750 aa, 1 stop
><MW: 84305, pI: 6.93, NX(S/T): 10
MWNLHETDSAVATARRPRWL CAGALVLAGGFLLGLFLGWFIKSSNEATNITPKHNMKAFL
DELKAENIKKFLHNFTQIPHLAGTEQNFQLAKQIQSQWKEFGLDSELAHYDVLLSYPNKTH
PNYISIIINEDGNEIFNTSLFEPPPPGYENVSDIVPPFSAFSPQGMPEGDLVYVNYARTEDFF
KLERDMKINCSCGKIVARIYRGVFRGNKVKNQLAGAKGVILYSDPADYFAPGVKSYPDGWNL
PGGGVQRGNILNLNGAGDPLTPGYPANFYAYRRGIAEAVGLPSIPVHPIGYYDAQKLEKMG
GSAPPDSSWRGSLKVPYNVGPFGFTGNFSTQKVKMHIHSTNEVTRIYNVIGTLRGAVEPDYV
ILGGHRDSWVFGIDPQSGAAVVHEIVRSFGTLKKEGWRPRRTILFASWDAEEFGLLGSTEW
AEENSRLQLQERGVAYINADSSIEGNYTLRVDCTPLMYSLVHNLTKELKSPDEGFEGKSLYES
WTKKSPSPFSGMPRIKSLGSGNDFEVFFQRLGIASGRARYTKNWETNKFSGYPLYHSVYET
YELVEKFYDPMFKYHLTVAQVRGGMVPELANSIVLPFDCRDYAVVLRKYADKIYSISMKHPQ
EMKTYSVSFDLSLFAVKNFTEIASKFSERLQDFDKSNPIVLRMMNDQLMFLERAFIDPLGLP
DRPFYRHVIYAPSSHNKYAGESFPGIYDALFDIESKVDPSKAWGEVQRQIYVAAFTVQAAAE
TLSEVA
```

Signal sequence:

amino acids 1-40

N-glycosylation sites.

amino acids 76-80, 121-125, 140-144, 153-157, 195-199, 336-340,
459-463, 476-480, 638-642

Tyrosine kinase phosphorylation sites.

amino acids 363-372, 605-613, 606-613, 617-626

N-myristoylation sites.

amino acids 85-91, 168-174, 252-258, 256-262, 282-288, 335-341,
360-366, 427-433, 529-535, 707-713